

**МАСТЕРА
ПСИХОЛОГИИ**

Скотт Миллер

ПСИХОЛОГИЯ РАЗВИТИЯ

методы исследования

второе международное издание



ПИТЕР®

МАСТЕРА ПСИХОЛОГИИ

Scott A. Miller

Developmental Research Methods

Second Edition

PRENTICE-HALL

Скотт Миллер

ПСИХОЛОГИЯ РАЗВИТИЯ: методы исследования

второе международное издание



Санкт-Петербург
Москва • Харьков • Минск

2002

Скотт Миллер

Психология развития: методы исследования

2-е международное издание

Серия «Мастера психологии»

Перевел с английского В. Белоусов

Главный редактор
Заведующий редакцией
Руководитель проекта
Ведущий редактор
Художественный редактор
Редактор
Корректоры
Верстка

Е. Строганова
Л. Винокуров
В. Попов
А. Борин
В. Шимкевич
В. Попов
Н. Баталова, М. Одинокова
Е. Кузьменок

ББК 88.8 УДК 159.92

Миллер С.

М60 Психология развития: методы исследования — СПб.: Питер, 2002. — 464 с.: ил. — (Серия «Мастера психологии»)

ISBN 5-318-00472-5

Трудность проведения исследования в области психологии развития — сквозная тема предлагаемой книги. Автор рассматривает ряд основных законов и принципов, которые необходимо учитывать при осуществлении собственных и оценке сторонних исследований. Это издание преследует три основные цели. Первая из них — помочь выработать навыки, необходимые для того, чтобы качественно изучать различные феномены в психологии развития, в связи с чем представлены разного рода принципы и инструкции, часть которых касается непосредственно вопросов развития, а другие относятся к психологии в целом. Вторая цель заключается в ознакомлении читателей с важнейшими направлениями научной работы в данной области. Третья цель — сформировать не только навыки критической оценки любого исследования, но и умение на его основе сделать адекватные выводы.

Эта книга адресована всем желающим углубить свои знания о проблемах и методах исследований в психологии развития: студентам, аспирантам, преподавателям факультетов психологии, психологам и психотерапевтам.

© 1998, 1987 by Prentice Hall, Inc.

© Перевод на русский язык, В. М. Белоусов, 2002

© Издательский дом «Питер», 2002

Права на издание получены по соглашению с Prentice Hall, Inc.

Все права защищены. Никакая часть данной книги не может быть воспроизведена в какой бы то ни было форме без письменного разрешения владельцев авторских прав.

ISBN 5-318-00472-5

ISBN 0-13-398892-9 (англ.)

ЗАО «Питер Бук». 196105, Санкт-Петербург, Благодатная ул., д. 67.

Лицензия ИД № 01940 от 05.06.00.

Налоговая льгота – общероссийский классификатор продукции ОК 005-93, том 2; 953000 – книги и брошюры.

Подписано в печать 27.12.01. Формат 70×100^{1/16}. Усл. п. л. 38,7. Тираж 5000 экз. Заказ № 2504.

Отпечатано с готовых диапозитивов в ФГУП «Печатный двор» им. А. М. Горького
Министерства РФ по делам печати, телерадиовещания и средств массовых коммуникаций.
197110, Санкт-Петербург, Чкаловский пр., 15.

Оглавление

Предисловие	9
Глава 1. Введение	11
Цели книги	13
Этапы исследовательской программы	14
Структура книги	19
Контент-анализ	20
Резюме	20
Упражнения	22
Глава 2. Общие принципы	23
Переменные	25
Валидность	26
Отбор	28
Контроль	30
Субъектные переменные	34
Переменные, поддающиеся и не поддающиеся манипуляциям	34
Возраст в качестве переменной	36
Результаты	37
Главный эффект	38
Взаимодействие	38
Причины искажения валидности	41
Резюме	42
Упражнения	44
Глава 3. План	45
Возрастное сравнение	46
Лонгитюдный метод	47
Метод поперечных срезов	52
Более сложные планы	55
Сравнение условий	60
Внутрисубъектные/межсубъектные планы	60
Планы уравненных групп	66
Планы временных серий	68
Корреляционное исследование	70
Корреляция и причинность	71
Способы повышения обоснованности выводов о причинно-следственных связях	72
Резюме	76
Упражнения	78
Глава 4. Измерение	80
Базовые понятия	80
Операционализация	81
Квантование	83
Аспекты измерения	84
Качество измерения	87
Тесты	89

Валидность	90
Надежность	93
Регрессия	94
Методы наблюдения	97
Что наблюдать	97
Как наблюдать	101
Определение точности результатов наблюдения	103
Резюме	107
Упражнения	108
Глава 5. Процедура	110
Стандартизация	110
Отклонения от стандартной процедуры	113
Сверхстандартизация	114
Некоторые систематические ошибки	116
Реактивность и связанные с ней проблемы	119
Установка на определенный ответ	121
Коммуникация между испытуемыми и диффузия	124
Предубеждения экспериментатора	126
Выбывание испытуемых	129
Резюме	130
Упражнения	132
Глава 6. Условия проведения исследования и его цели.	133
Условия проведения исследования	134
План 1: «лаборатория—лаборатория»	135
План 4: «полевые условия—полевые условия»	137
План 2: «лаборатория—полевые условия»	140
План 3: «полевые условия—лаборатория»	141
Анализ и оценка	141
Экологические системы	144
Прикладные исследования	149
Оценка	150
Вмешательство	154
Резюме	157
Упражнения	159
Глава 7. Статистика	160
Использование статистических показателей	160
Дескриптивные статистические показатели	161
Статистические показатели, выводимые логическим путем	163
Выбор статистического показателя	168
Уровень измерения	168
Распределение показателей	168
План исследования	171
Меры связи	175
Альтернативы и более сложные варианты дисперсионного анализа	178
Плановые сравнения	178
Величина эффекта	180
Мультивариантный дисперсионный анализ	181

Множественный регрессионный анализ	183
Общие принципы	185
Резюме	186
Упражнения	188
Глава 8. Этика	189
Нормы и процедуры	190
Права испытуемых	194
Информированное согласие	194
Ненанесение ущерба	200
Конфиденциальность	204
Резюме	206
Упражнения	208
Глава 9. Публикация	209
Общие замечания	210
Разделы статьи	213
Введение	213
Метод	214
Результаты	215
Обсуждение	217
Другие разделы	218
Частные аспекты стилистики	220
Нормы АПА	220
Общие вопросы английского языка	222
Резюме	224
Упражнения	225
Глава 10. Младенчество	226
Общие вопросы	226
Отбор	226
Значение состояния младенца	228
Оценка реагирования	231
Возрастные сравнения	232
Восприятие в младенчестве	234
Метод предпочтений	236
Габитуация/дегабитуация	239
Обусловливание	241
Визуальный обрыв	243
Когнитивное развитие в младенчестве	245
Социальное развитие в младенчестве	253
Привязанность	254
Темперамент	261
Резюме	267
Упражнения	269
Глава 11. Когнитивное развитие	271
Подход Пиаже	271
Исследования Пиаже	271
Спорные вопросы и дальнейшие исследования	278

Подход, основанный на тестировании интеллекта	286
Сущность <i>IQ</i> -тестов	286
Примеры тестов	288
Спорные вопросы и исследовательские парадигмы	292
Память	296
Память в младенчестве	296
Мнемонические стратегии	299
Конструктивная память	304
Теория психики	306
Методы изучения	306
Спорные вопросы исследований	310
Резюме	313
Упражнения	315
Глава 12. Социальное развитие	317
Нравственное развитие	317
Поведенческий аспект	318
Эмоциональный аспект	325
Когнитивный аспект	330
Я-концепция	341
Самосознание	341
Самооценка	344
Половые различия	348
Общие замечания	348
Методики изучения половых различий	351
Социализация	360
Лабораторные исследования	360
Натуралистическое наблюдение	364
Методики вербального отчета	368
Резюме	373
Упражнения	376
Глава 13. Старение	378
Общие замечания (на примере исследования <i>IQ</i>)	379
Отбор	379
Измерение	387
План исследования	394
Память	398
Узнавание и воспроизведение	399
Мнемонические стратегии и конструктивная память	401
«Повседневная» память и долговременная память	403
Личность и социальное развитие	410
Удовлетворенность жизнью	410
Методики изучения личностных особенностей	413
Общие замечания	416
Резюме	418
Упражнения	420
Приложение	421
Литература	422
Глоссарий	455

Предисловие

Эта книга адресована всем, кто хочет углубить свои знания о проведении исследований в области психологии развития. Она не научит вас всему, что связано с процессом исследования, — ни одна книга не может этого сделать. Но я надеюсь, что в ней вы найдете ценную основу, ряд общих законов и принципов, которые помогут как в осуществлении собственных исследований, так и при оценке исследований, проведенных другими.

Я попытался найти равновесие между общим и частным. Этот баланс отражен в структуре книги: за первыми 9 главами, в которых рассматриваются общие вопросы методологии исследования, следуют 4 (несколько более длинные) главы, посвященные конкретным темам исследования в психологии развития. Попытка соблюсти равновесие также отражена в подходе к рассмотрению самого исследования. С одной стороны, это и не абстрактный трактат об «исследовательском анализе», а с другой — и не практическое руководство. В книге изложено то, с чем, по моему мнению, сталкивается большинство из нас при проведении исследования — в ней обращается внимание на моменты, которые необходимо учесть, на решения, которые придется принять, на затруднения, которые нужно преодолеть на каждом из этапов подготовки и проведения исследования. Я надеюсь, что в этой книге мне удалось передать состояние творческого волнения, которое охватывает исследователя в процессе изучения действительно важных вопросов.

Основными читателями этой книги, конечно, станут те, кто работает в исследовательских лабораториях или слушает курс лекций о методах исследования в психологии развития или в детской психологии. Я полагаюсь на то, что каждый, кто слушает эти курсы, уже прослушал предварительный курс психологии развития или детской психологии. Знание других предметов (например, статистики, методов психологического исследования) будет ценным, но оно необязательно. Кроме того, преподаватели могут приспособить эту книгу для студентов последних курсов и аспирантов.

Это издание книги несколько отличается от предыдущего. Некоторые изменения коснулись структуры книги. Главы, посвященные специальным темам, из середины книги перенесены в конец — большинство из нас (по крайней мере я) излагают этот курс именно в такой последовательности. Изменилась также структура глав, в которых характеризуются базовые понятия; я надеюсь, что новая компоновка более логична. Конечно, преподаватель всегда может излагать содержание глав — или их отдельных разделов — в удобном для себя порядке. В это издание был добавлен ряд новых тем; среди них — прикладное исследование, планы временных серий, экологический системный подход Бронфенбреннера, Я-концепция и теоретические вопросы психологии. Даже там, где название тем осталось прежним, содержание было обновлено; более половины имеющихся в этой книге ссылок отсутствуют в первом издании. Наконец, было включено два вспомогательных раздела: список основных терминов и упражнения в конце каждой главы. Предлагаемые упражнения разнообразны по форме и по целям. В зависимости от предпочтений преподавателя и доступных ресурсов, каждое конкретное упражнение может стать основой для практического исследования (проводимого либо индивидуально, либо в группе), группового обсуждения или просто осуществляемого самостоятельно мысленного эксперимента.

Я выражаю благодарность многим людям, оказавшим помощь при работе над обоими изданиями книги. Общую поддержку оказало отделение психологии университета штата Флорида. Джон Айли и Сьюзан Уиллиг (John Isley, Susan Willig) оставались научными редакторами издательства «Прентис-Холл» на протяжении неожиданно затянувшегося периода «вынашивания» книги в ее первом издании, а Пит Джэнзоу, Николь Синьоретти и Айлин Кэлиш (Pete Janzow, Nicole Signoretti, Ilene Kalish) взяли на себя этот труд на время довольно длительной подготовки пересмотренного варианта. Я признателен всему штату работников издательства «Prentice Hall» за проявленные компетентность и терпение. Оба издания книги выиграли благодаря участию группы замечательных рецензентов. Рецензентами первого издания были Дэниэл Б. Кайе, Марша Лисс, Мэрион Перлмуттер, Майкл Прессли и Патрисия Уорден (Daniel B.Kaye, Marsha Liss, Marion Perlmutter, Michael Pressley, Patricia Worden), а Барбара Майерс и Адриен Томер (Barbara Myers, Adrien Tomer) были рецензентами второго варианта. Я также благодарю Карла Фон Байера (Carl Von Bayer) за комментарии ко второму изданию и Джеймса Элджину (James Algina) за рекомендации к главе по статистике. Наконец, я благодарю Хизер Холмс (Heather Holmes) за ее всестороннюю помощь, включая подготовку контент-анализа.

Глава 1

Введение

Эгоцентризм маленьких детей всегда вызывал у психологов огромный интерес. Эгоцентризм — это неспособность отстраниться от собственной точки зрения и принять во внимание точку зрения другого человека. Старшим детям и взрослым кажется естественным, что другие воспринимают окружающее не так, как они, видят не то, что они, знают не то, что знают они, желают не того, что желают они. Однако этот факт не всегда очевиден для маленьких детей. Они зачастую ведут себя так, будто все разделяют их точку зрения на мир, отсюда и термин «эгоцентричный».

Как, например, ведет себя маленький мальчик в экспериментальной ситуации, когда его просят вообразить, что он покупает своей матери подарок на день рождения (Flavell, Botkin, Fry, Wright, & Jarvis, 1968)? Перед ребенком выкладывают ряд предметов, уместных в качестве подарка для лиц разного пола и возраста. Направляется ли ребенок незамедлительно к шелковым чулкам или взрослым книгам? Хотя такая реакция и возможна, для 3–4-летних детей она нехарактерна. Скорее всего, выбор падет на новенький игрушечный грузовик. Малыш знает, чего он хочет, как же может мать не хотеть того же?

В некотором смысле авторы учебников сильно напоминают этого маленького ребенка, стоящего перед столом с подарками. Для них занимательность и ценность, даже прелесть их предмета самоочевидна. Реакцией на просьбу объяснить, почему кого-то еще должна волновать эта тема, вероятно будет удивление или обида; как кто-то может не понимать, что это удивительный и жизненно важный предмет? Чем не новенький грузовик?

Тем не менее (и тут проявляется эгоцентризм) вряд ли можно утверждать, что интерес к исследованиям в области психологии развития требует каких-либо обоснований. Что может быть более естественным, чем необходимость изучать процесс человеческого развития? Если этого объяснения все-таки недостаточно, можно легко привести ряд доказательств. Несомненно, ни одна отрасль психологии не является столь обширной, как психология развития. И несомненно, ни одна от-

расль психологии не рассматривает больше фундаментальных научных вопросов, чем психология развития, поскольку последняя включает в себя все темы (восприятие, мышление, личность и т. д.) и добавляет к ним один-единственный основополагающий вопрос: как люди становятся такими, какие они есть? Каким образом, к примеру, происходит так, что практически все люди в конце концов начинают пользоваться сложнейшей языковой системой? Откуда проистекают индивидуальные различия в интеллекте или характере? Как сказывается воспитание в раннем возрасте на дальнейшем развитии? Подобного рода вопросы открывают сущность того, о чем потенциально как наука может поведать нам психология.

Эти вопросы представляют не только научный интерес. Ни в одной другой области рассматриваемые проблемы не являются столь важными для каждого человека. Вернемся к некоторым из вопросов, обозначенных в предыдущем абзаце. Проблема психического опыта первых лет жизни и последующего развития может быть чрезвычайно интересна для исследователя; однако это же и насущный практический вопрос для любого родителя, заботящегося об оптимальном развитии своего ребенка. То, что люди различаются по уровню интеллекта, влечет за собой ряд в высшей степени важных теоретических вопросов, но этот факт имеет также громадное социальное значение. Одной из данностей специализации в области психологии развития является ощущение действительной значимости изучаемых вопросов.

Однако найти ответы на эти вопросы непросто. Фактически самыми сложными из них оказываются наиболее существенные, фундаментальные. Трудность проведения добротного исследования — сквозная тема всей этой книги, и поэтому здесь ее можно не обсуждать. Но давайте, прежде чем переходить к общим моментам, кратко рассмотрим пример. Это проблема, которую мы затрагивали уже дважды: установление связи между методами воспитания, которых придерживаются родители, и развитием ребенка. Каким должен быть подход к изучению данной проблемы, чтобы быть научным?

Для каждого, имеющего хотя бы элементарную подготовку в области методологии, общий ответ на этот вопрос очевиден: с помощью контролируемого экспериментального исследования (если этот ответ не очевиден, он станет таковым после прочтения главы 2). Можно, например, в произвольном порядке распределить новорожденных по разным семьям, где родители придерживаются разных взглядов на проблему воспитания. Тогда эффект воспитания можно было бы отделить от влияния на развитие ребенка родительских генов. Или же исследователь может решить произвольно предписать разным семьям разные способы воспитания. Эта процедура дала бы возможность избежать непредсказуемости выбора родителями того или иного стиля воспитания и получить ясное представление о роли методов воспитания самих по себе. Исследователь может даже в целях сравнения подобрать группу родителей, воспитывающих своих детей в соответствии с собственными представлениями. В любом случае за детьми наблюдали бы в процессе их взросления, а для оценки особенностей их развития производили бы тщательные измерения. Если такое исследование можно было бы осуществлять хотя бы в течение нескольких лет, мы знали бы гораздо больше об эффектах разных стилей воспитания.

Нет необходимости говорить, что описанная выше исследовательская программа — нечто из области научной фантастики (или учебников), но никак не реально-сти. Мы не продлеваем подобного рода экспериментов и, надеюсь, никогда не будем проводить¹. Этических проблем здесь более чем достаточно, чтобы запретить и такое исследование. Но даже если бы их не было, практические трудности самого проведения исследования оказались бы почти непреодолимыми. Оба этих обстоятельства — этические барьеры и практические ограничения — исключают возможность проведения множества «безупречных экспериментов», которые с легкостью мог бы предложить любой психолог. В результате для сбора необходимой информации мы вынуждены прибегать к методам, не отвечающим всем критериям научности. То, что такие методы действительно существуют, и то, что благодаря им можно получить действительно достоверные результаты, — еще одна сквозная тема этой книги. Однако выбрать подходящий метод и получить нужную информацию зачастую также непросто.

Подведем итог вышесказанному. Психология развития обращается к вопросам, имеющим как научную, так и практическую ценность. Изучение этих вопросов нередко связано с серьезными трудностями, которые неизбежно ограничивают возможности познания. Таким образом, перед нами сфера научной деятельности, результаты которой потенциально имеют огромное значение, где проблемы, стоящие перед исследователем — крайне сложны, накопление фактов происходит медленно, но эти факты обладают ценностью, иными словами — это идеальная сфера деятельности для энергичного и предприимчивого исследователя.

Цели книги

Эта книга преследует три основные цели. Первая и наиболее очевидная из них — помочь сформировать навыки, необходимые для грамотного проведения исследования в области психологии развития. Для этого в книге представлены разного рода теоретические принципы и практические рекомендации. Часть этих принципов касается непосредственно вопросов развития; другие относятся к психологии в целом. Фактически некоторые из них имеют отношение даже не к самой психологии, а к общенаучной методологии. Однако там, где это возможно, я веду рассуждение в рамках психологии развития. А как уже говорилось, в самой психологии развития достаточно проблем, чтобы озадачить любого исследователя.

Вторая цель заключается в ознакомлении читателей с важнейшими направлениями исследований в данной области. В конечном счете никто не «изучает развитие вообще», исследования всегда проводятся в конкретных областях, в каждой из кото-

¹ Есть исторические свидетельства, возможно недостоверные, о правителях, которые довольно последовательно проводили эксперименты по воспитанию детей. Один из них — правивший в XIII веке король Фридрих II. Его эксперимент, как и многие другие соответствовавший представлениям его времени, дал интересные результаты, однако совсем не те, на которые рассчитывал Фридрих: он приказал кормилицам и няням кормить, купать и мыть детей, однако ни в коем случае не откликаться на их лепет и не разговаривать с ними, поскольку он хотел узнать, заговорят ли дети на древнееврейском, греческом, латыни, арабском или, возможно, на языке родителей, давших им жизнь. Но все его труды пропали даром, поскольку, не ощущая ласки, не видя улыбающихся лиц и не слыша нежных слов своих кормилиц, все дети умерли (Ross & McLaughlin, 1949, p. 366).

рых есть свои методологические трудности. В одной книге нельзя охватить все интересные темы или полностью раскрыть какую-то одну тему. Однако можно заложить основы для понимания наиболее интересных и хорошо изученных вопросов.

Третья цель — сформировать навыки, необходимые для того, чтобы уметь критически оценивать любое исследование и делать на его основе собственные выводы. Эти навыки, конечно, неотделимы от навыков проведения самого исследования, однако большинство из нас, скорее всего, будет использовать их значительно чаще. Не каждый собирается проводить исследование в области психологии развития, но каждый является потребителем результатов подобных исследований. Вернемся вновь к некоторым практическим вопросам. Является ли в некоторых случаях оправданным физическое наказание или такие дисциплинарные методы абсолютно неприемлемы? Провоцирует ли насилие на экране детскую агрессию? Стоит ли разрабатывать программы раннего развития для так называемых «неблагополучных» детей? Обоснована ли вообще отправка на пенсию по возрасту, и если да, то при какого рода занятиях? Подобные вопросы волнуют каждого родителя, каждого налогоплательщика, каждого избирателя. Более разумные ответы даст тот, кто имеет представление о выводах, сделанных на основе соответствующих исследований. Еще более разумные ответы даст тот, кто имеет представление о методах, стоящих за исследованиями, и может, анализируя эти выводы, здраво оценить все их сильные и слабые стороны, а также тому, насколько ему можно доверять.

Этапы исследовательской программы

Как спланировать исследование для того, чтобы его конечным продуктом явилась ценная с научной точки зрения информация? Большую часть этих моментов мы детально рассмотрим в следующих разделах. Целью же этого раздела является предварительное ознакомление с тем, что необходимо для проведения добротного исследования, — беглый обзор вопросов, которые будут освещены в дальнейшем.

Стартовой точкой любой успешной исследовательской программы являются *хорошие идеи*. Это одновременно и наиболее очевидное из требований, и требование, выполнению которого труднее всего научить. Поскольку оно и очевидно, и плохо поддается разъяснению, обычно при обсуждении вопроса о том, как проводить исследование, критерием достоинств идеи пренебрегают, при этом все внимание сосредоточивается на навыках, необходимых для проверки этих идей, какими бы они ни были. Это справедливо и для нашей книги. Однако важно помнить, что ни один технический навык не спасет исследование, если гипотезы, положенные в его основу, непродуктивны. Важно также осознавать, что исследователь, открытие которого становится подлинным событием в мире науки, отличается от рядового не столько техникой проведения исследования, сколько степенью оригинальности и проницательности при рассмотрении определенного вопроса.

Второй критерий — *осведомленность об уже проделанной работе*. Каждый, кто начинает осуществление исследовательской программы, должен обладать исчерпывающей информацией о том, что сделано по интересующей его теме. Фактически, если следовать логике, этот этап следовало бы поставить на первое место, поскольку без знаний о том, что было сделано прежде, в голове не могут родиться действительно хорошие идеи. В любом случае знание литературы является суще-

ственным моментом, когда исследователь начинает размышлять о ценности собственных идей. Практически нет смысла реализовывать в жизнь блестящую исследовательскую идею, если кто-то ее уже реализовал. Чаще даже случается так, что если бы исследователь знал об уже проделанной работе, то принял бы иное решение относительно важнейших моментов процедуры. Самое неприятное для исследователя — приложить массу усилий для проведения исследования и только затем узнать, что подобное исследование уже проводилось и все его усилия были напрасны. Следить за литературой — непростая задача в то время, когда ежегодно специальные журналы публикуют тысячи статей по психологии развития. К счастью, есть вспомогательные издания. Уже давно выходят реферативные журналы, единственной целью которых является краткий обзор статей и книг по психологии. В последние годы некоторые из этих реферативных сборников появились в компьютерной сети — эта технология может существенно ускорить поиск нужной литературы. В табл. 1.1 кратко охарактеризованы реферативные источники, наиболее полезные для специализирующихся в области психологии развития. Ценную помощь в поиске литературы оказывают также различные книги и журналы, посвященные исключительно обзору статей по основным темам. Некоторые из наиболее полезных перечислены и кратко охарактеризованы в табл. 1.2. В таблицу также включены некоторые из основных профессиональных журналов, в которых публикуются исследования в области психологии развития. Прежде чем принять окончательное решение по практическим вопросам, полезно просмотреть последние номера этих журналов. Наконец, можно получить ясное представление о проделанной работе не из письменных источников, а проконсультировавшись у квалифицированного исследователя. Помимо обращения к библиографическим источникам полезно обсуждать свои идеи со специалистами, что явится ценным элементом процесса разрешения многих спорных вопросов.

Таблица 1.1

Реферативные источники в области психологии

Источник	Формат	Охватываемый период	Тематика
«Child Development Abstracts and Bibliography»	Печатный	С 1927 г. до настоящего времени	Психология развития
«PsyScan: Developmental Psychology»	Печатный	С 1980 г. до настоящего времени	Психология развития
«Psychological Abstracts»	Печатный	С 1927 г. до настоящего времени	Психология в целом
PsychINFO	On-line	С 1921 г. до настоящего времени	Психология в целом
PsychLIT	CD-Rom	С 1974 г. до настоящего времени	Психология в целом

**Источники обзорных статей и отчетов об исследованиях
в области психологии развития**

Обзорная литература	«Толстые» журналы
Advances in Child Development and Behavior (первый том вышел в 1963 году, с тех пор выходит практически ежегодно) Annals of Child Development (выходит ежегодно с 1984 года) Annual Review of Psychology (выходит ежегодно с 1950 года) Handbook of Child Psychology, 5-е издание, 1998 Handbook of the Psychology of Aging, 4-е издание, 1996 Minnesota Symposia on Child Psychology (выходит ежегодно с 1967 года) New Directions for Child Development (многотомник, выходящий с 1978 года) Developmental Review (реферативный журнал исследований в области психологии развития) Psychological Bulletin (реферативный журнал по всем областям психологии)	British Journal of Developmental Psychology Child Development Cognitive Development Developmental Psychology Experimental Aging Research Genetic Psychology Monographs Human Development Infant Behavior and Development International Journal of Behavioural Development Journal of Applied Developmental Psychology Journal of Experimental Child Psychology Journal of Genetic Psychology Journal of Gerontology Merrill-Palmer Quarterly Monographs of the Society for Research in Child Development Psychology and Aging Social Development

За формулированием исследовательских гипотез идет этап их перевода в *план эксперимента*. Как уже упоминалось выше, технически совершенный план не имеет смысла, если выдвинутые идеи не стоят того, чтобы их проверять. Сейчас я добавлю к этому нечто противоположное: любая, даже самая блестящая гипотеза превращается в ничто, если ей нельзя придать форму, в которой ее можно научно проверить. Планирование эксперимента — центральная тема последующих глав. Пока же сделаем лишь два замечания. Первое повторяет уже сказанное. В психологии развития очень часто случается так, что этические и практические барьеры препятствуют проведению исследований, которые, с чисто научной точки зрения, были бы идеальными для изучения определенного вопроса. Трудность заключается в том, чтобы спланировать альтернативные процедуры, позволяющие прийти к валидным заключениям. Второй момент: планирование исследований в области психологии развития зачастую осложняется тем фактом, что в качестве исследуемой переменной выступает возраст. Как мы увидим в дальнейшем, возраст — это переменная, с которой в определенном отношении особенно трудно работать. Однако, несомненно, возрастные изменения представляют огромный интерес для большинства специалистов в области психологии развития.

Теперь наш воображаемый исследователь находится на этапе, когда он уже сформулировал идею, просмотрел нужную литературу и принял (по крайней мере предварительное) решение относительно плана эксперимента. Следующий шаг — *получение общественного одобрения* — то есть представление научного предложения на рассмотрение университетского комитета, контролирующего этичность проводимых исследований. Этике посвящена глава 8, в ней будет много сказано и о процедурах, которым нужно следовать, и о критериях этики, которым нужно соответствовать. Пока же я ограничусь одним существенным моментом, который касается необходимости оценки этичности исследования. Конечно, исследователи должны делать все возможное для обеспечения этичности собственных исследовательских проектов. Однако они не должны принимать это решение в одиночку; исследование может проводиться только после того, как независимая комиссия убедится в том, что оно отвечает требованиям этики.

Хотя следующий шаг необходим не во всех исследовательских проектах, в определенных случаях он оправдан. Это *пилотажное исследование* — то есть предварительное тестирование до начала самого эксперимента. Для использования пилотажной проверки есть два существенных основания, оба они приобретают особое значение при работе с детьми. Во-первых, экспериментатору необходимо получить практические навыки проведения определенных процедур и в работе с испытуемыми, дабы свести к минимуму возможность ошибки в реальных условиях эксперимента. Во-вторых, необходимо проверить все то, что может помешать процессу исследования. Ясны ли инструкции? Достаточно ли времени отводится на тестирование? Убедительно ли выглядит обманный прием, задуманный экспериментатором? В разных проектах будут актуальны разные проблемы, однако общий вопрос останется неизменным: все ли готово для начала исследования?

Если допустить, что ответ на этот вопрос положителен, то следующий шаг вновь будет очевиден: *подбор испытуемых*. Хотя этот шаг кажется естественным, во многих учебниках по методологии он опускается, и план эксперимента каким-то магическим путем превращается в эмпирические данные, без хлопотного промежуточного этапа поиска людей, при участии которых этот план будет реализован. В действительности, многие исследователи значительную часть своей карьеры тратят не на интересный процесс обдумывания исследования, а на гораздо более утомительный труд поиска испытуемых. Это особенно верно в отношении специалистов в области психологии развития, у которых нет готовой группы испытуемых. Исследователь, который хочет изучать период младенчества, не может просто вывесить объявление, приглашающее малышей принять участие в эксперименте; он должен каким-то образом найти семьи с младенцами и убедить родителей принести своих малышей на обследование.¹ Исследователь, намеревающийся изучать большие выборки 5-, 7- и 9-летних детей, в поисках достаточного количества испытуемых будет вынужден обратиться в школы. Тем, кто хочет изучить возрастные изменения в старости, потребуются найти и привлечь к исследованию пожилых людей, возможно, заручившись поддержкой различных организаций, обслуживающих людей преклонного возраста. Подбор всех этих испытуемых может оказаться особенно проблематичным.

Трудно предложить конкретные правила привлечения испытуемых, поскольку в разных местах существуют свои особенности. Однако можно дать несколько общих рекомендаций. Во-первых, нужно иметь запас времени. Практически всегда на исследование уходит больше времени, чем ожидает начинающий исследователь, и типичной причиной срыва сроков является трудность в привлечении участников. Вторая рекомендация — уметь убеждать, представляя свои исследовательские предложения тем, кто должен принять решение об участии (директорам школ, учителям, родителям, самим детям). Как подчеркивается в главе 8, посвященной этике, первое, о чем нужно помнить при ознакомлении потенциальных испытуемых с исследованием, — о собственной честности и открытости, так чтобы согласие на участие не было получено обманным путем. Однако не менее важно разъяснить ценность исследования, иначе, возможно, никто не захочет принять в нем участие. Наконец, вероятно наиболее плодотворный путь — найти исследователя, имеющего опыт в изучении данной популяции, и попросить у него или у нее совета о том, как лучше действовать.

Заметьте, что основная проблема, связанная с привлечением испытуемых, заключается не в неудобствах, с которыми сталкивается экспериментатор, и не в потере времени. Проблема состоит в том, что одним из условий грамотного проведения исследования является подбор определенного рода испытуемых. Как мы увидим в дальнейшем, даже добросовестно проведенный эксперимент не имеет смысла, если экспериментатору не удалось набрать именно таких участников, каких нужно.

Как только сформирована группа испытуемых, можно начинать обследование. На этом этапе приобретают значимость *исследовательские навыки экспериментатора*. Употребленное здесь выражение «исследовательские навыки» означает весь набор способностей, необходимых для реальной работы с испытуемыми, будь то непосредственное взаимодействие, наблюдение или оценка поведения. Таким образом, имеют значение вопросы такого рода: передают ли инструкции достаточно четко что необходимо? Не влияет ли сам экспериментатор на получаемые им результаты выражением лица или непреднамеренным подкреплением? Аккуратно ли регистрируются ответы? Иначе говоря, вопрос состоит в том, может ли заблаговременно разработанное на бумаге исследование быть адекватно воспроизведено в экспериментальных условиях? И вновь ясно, что успешное прохождение предыдущих этапов исследовательской программы будет бессмысленным, если не удастся добиться успешного преодоления этого этапа. Исследователь, к примеру, составил великолепный план изучения процесса решения задач 5-летними детьми, однако результаты вряд ли окажутся значимыми, если он не имеет понятия о том, как общаться с 5-летками, и поэтому вызывает у них или испуг, или изумление.

Тема исследовательских навыков затрагивается в разных разделах этой книги, однако более подробно она рассмотрена в главе 5. Как уже указывалось, некоторые моменты являются общими, имеющими отношение к психологии в целом, другие — касаются непосредственно психологии развития. Хотя трудности могут возникать в любом исследовании, изучение вопросов развития часто связано со специфическими проблемами, проистекающими из специфичности обследуемых групп. Навыки, достаточные для тестирования студента колледжа, могут оказывать

ся недостаточными для работы с плачущим младенцем, застенчивым дошкольником или недоверчивым 80-летним стариком. Ситуация осложняется еще более, когда в один исследовательский план включено сразу несколько возрастных групп.

Ни об одном аспекте методологии исследования нельзя надлежащим образом узнать лишь из учебника. Однако ни один учебник не может дать практических навыков непосредственной работы с испытуемыми. Хотя ряд правил можно изложить в словесной форме, единственно реальный способ приобрести опыт в работе с младенцами, дошкольниками или пожилыми людьми — проводить значительное время, работая с ними на практике.

Завершение процесса обследования не означает окончания исследования. Следующий этап — *статистический анализ данных*. Теперь вопрос, на который нужно ответить, состоит в том, связаны или нет разнообразные исследованные факторы с какими-либо устойчивыми, значимыми паттернами результатов. В подавляющем большинстве исследований на этот вопрос помогают ответить определенные, тщательно разработанные статистические процедуры. Статистический анализ сам по себе не ответит на вопросы о теоретическом и практическом смысле результатов, но тем не менее поможет установить пределы для их интерпретации.

Статистический анализ — это серьезная тема, предмет отдельных учебных курсов и книг. Эта книга не претендует на его освещение. Однако в главе 7 можно найти краткое изложение некоторых общих принципов статистики.

Заключительной фазой исследования является *сообщение* о том, что было сделано и что было обнаружено. Наука — это поле общей информации, а научное открытие — не открытие, пока о нем не узнают другие. Традиционный способ сообщения о результатах в области психологии развития — публикация их в специальном журнале. Для этого исследователь должен подготовить краткий и ясный отчет об исследовании. Рекомендации по подготовке таких отчетов изложены в главе 9.

Структура книги

Следующие пять глав касаются общих принципов планирования и процедуры эксперимента. Фактически глава 2 называется «Общие принципы»; в ней рассмотрены такие базовые понятия, как независимые и зависимые переменные, экспериментальный контроль и разные формы валидности. Глава 3 называется «План»; в ней обсуждаются способы построения исследования, а также сравнения и выводы, возможные при разных подходах. Поскольку книга написана с позиции психологии развития, особое внимание уделено планированию эксперимента с целью сравнения разных возрастных групп. В главе 4 «Измерение» акцент с независимой смещается на зависимую переменную; при этом рассматриваются способы измерения того, что мы получаем на выходе. Решения как о плане, так и процедуре эксперимента нужно претворить в жизнь, и это является темой главы 5 «Процедура»: проблемы, которые могут возникнуть в процессе воплощения плана в реальных условиях, а также способы их преодоления. Наконец, глава 6 «Условия и цели» посвящена вариантам экспериментальной обстановки (например, структурированные лабораторные или естественные «полевые» условия), а также их плюсам и минусам.

сам. Кроме того, рассматриваются цели исследований в области психологии развития — близкая к предыдущей тема, поскольку цели являются одной из детерминант выбора экспериментальных условий.

В следующих трех главах описываются три важнейших этапа проведения исследования. В главе 7 представлены некоторые общие принципы статистической проверки и обоснования. Глава 8 посвящена этическим вопросам исследования в области психологии развития, а в главе 9 изложены рекомендации по написанию отчета о психологическом исследовании.

Заключительный раздел книги посвящен специальным областям исследования в психологии развития. В главе 10 рассматриваются методы изучения развития в младенчестве. Следующие две главы различаются по тематике в рамках одной возрастной группы. В главе 11 внимание сосредоточивается на способах изучения когнитивного развития, особенно в период раннего и среднего детства; в главе 12 речь идет об изучении социального развития. Наконец, глава 13 возвращает нас к хронологическим вопросам и посвящена методам изучения развития в пожилом возрасте.

Контент-анализ

Следует упомянуть еще одну особенность данной книги. В ней анализируется ряд параметров, по которым различаются исследовательские проекты и отчеты. Например, проводится исследование в естественных или лабораторных условиях, каков план эксперимента для сравнения разных возрастных групп? Многое будет сказано о том, какие можно и какие желательно принять решения относительно процедуры. Однако помимо возможного и желательного важно знать и фактически существующее — что делается на практике. Какие методологические решения принимают сегодня исследователи в области психологии развития? Рассмотрение актуальной исследовательской практики поможет придать конкретные формы тому, что иначе казалось бы довольно абстрактным изложением общих моментов.

С этой целью был проведен контент-анализ двух ведущих в данной области исследовательских журналов — «Child Development» и «Developmental Psychology». Обзор был основан на случайной выборке исследовательских отчетов, опубликованных в этих журналах в 1995 году — всего 74 статьи и 92 исследования. Параметры, которые при этом оценивались, перечислены и кратко охарактеризованы в табл. 1.3. Значение некоторых терминов может быть не совсем понятным; оно прояснится в ходе ознакомления с текстом. По мере изложения соответствующих тем будут представляться и результаты контент-анализа.

Резюме

Глава начинается с рассмотрения как значения исследований в области психологии развития, так и связанных с их проведением трудностей. Затем следует обсуждение трех основных целей книги: сформировать навыки, необходимые для проведения исследования психического развития человека, ознакомить с наиболее

Таблица 1.3

**Параметры, оценивавшиеся в контент-анализе публикаций
в области психологии развития**

Параметр	Описание
Количество возрастных групп	Количество отобранных и сравнивавшихся возрастных групп
Информация о выборке	Информация о таких характеристиках, как пол, раса, социальный статус
Информация об экспериментаторе	Информация о количестве экспериментаторов и их характеристиках (пол, возраст)
Отбор испытуемых	Каким образом были привлечены испытуемые и процент согласившихся на участие от первичного контингента
Выбывание испытуемых	Количество испытуемых, выбывших из исследования в ходе его проведения, и причины выбывания
Способ сравнения возрастных групп	Метод, которым сравнивались возрастные группы (например, метод поперечных срезов, лонгитюдный)
Надежность	Подсчитывался ли уровень согласованности оценок экспертов и наблюдателей
«Слепое» тестирование	В определенных случаях, неинформированность проводящих тестирование или наблюдение об испытуемых или гипотезах
Условия	Условия (например, лабораторные, естественные), в которых производились манипуляции с переменными, и их оценка
Экологический уровень	Экологическая система или экологические системы (например, семья, культура в целом), которые охватывает исследование

интересными и значимыми темами исследований в данной области, а также способствовать формированию навыков критической оценки этих исследований.

В среднем разделе главы перечислены этапы, которые нужно пройти для получения фактических результатов. Начальный этап имеет наибольшее значение, но его преодолению труднее всего обучить — это формулирование *хороших идей*, заслуживающих проведения эмпирического исследования. Тесно связанным с предыдущим и, вероятно, даже предшествующим этапом должно стать *ознакомление с работой, уже сделанной ранее*, поскольку любое исследование рождается из того, что было сделано прежде. Хорошие идеи должны воплотиться в *план проведения эксперимента*, на основе которого могут быть сделаны четкие и валидные выводы. До начала осуществления исследования необходимо получить *одобрение независимой комиссии*, и иногда бывает полезно провести *пилотажное исследование* с целью уточнения сомнительных моментов в процедуре и оттачивания исследовательских навыков. Еще одним существенным и зачастую трудным подготовительным этапом является *подбор испытуемых*: необходимо найти соответствующих опреде-

ленным требованиям потенциальных испытуемых и заручиться их согласием на участие. С началом исследования особое значение приобретают *исследовательские навыки* — то есть все те навыки, которые необходимы для взаимодействия с испытуемыми и беспристрастного наблюдения. За сбором данных следует их *статистический анализ*, с помощью которого выявляют достоверные и потенциально информативные паттерны результатов. Заключительный этап — *сообщение* о полученных результатах, обычно в форме публикации в специальном журнале.

Беглое рассмотрение этапов исследовательской программы служит для ознакомления с темами, которые будут подробно обсуждаться в дальнейшем. В завершающей части введения дано предварительное описание тем последующих глав. Кроме того, сделаны предварительные замечания о проведенном контент-анализе публикаций ведущих журналов по психологии развития, результаты которого включены в разные главы этой книги.

Упражнения

Изложенное выше наводит на мысль о том, что исследования в области психологии развития затрагивают ряд насущных практических вопросов личностного, социального, политического значения. Поразмышляйте об этом в течение недели, когда будете читать газеты или слушать новости. Для каких тем, о которых вы узнали из новостей, оказалось бы ценным понимание принципов психологии развития?

Один из способов получить общее впечатление об актуальных вопросах в определенной области исследований — просматривать последние выпуски ведущих в данной области научных журналов. Публикации в этих журналах, кроме того, могут подать вам идею собственного исследования. Выберите по меньшей мере три журнала из представленных в табл. 1.2 и найдите в библиотеке их последние тома (имейте в виду, что годовой том обычно состоит из нескольких отдельных номеров). Прочтите все заголовки и резюме наиболее заинтересовавших вас статей.

Выберите две из наиболее интересных для вас тем в области психологии развития и найдите литературу по ним с помощью PsychLit или PsychInfo.

Общие принципы

Знакомство с несколькими конкретными исследованиями поможет уяснить суть последующего материала. Поэтому ниже приведены два примера исследований в области психологии развития. Оба описания представлены в несколько упрощенном виде с тем, чтобы облегчить уяснение основных моментов.

Дюфресн и Кобасигава (Dufresne & Kobasigawa, 1989) пытались выявить детерминанты развития памяти в детском возрасте. Почему старшие дети запоминают лучше младших? Конкретная детерминанта, на которой сосредоточивалось исследование, обозначена как *время, отпущенное на запоминание*. Вопрос стоит следующим образом: что делают дети, получая возможность изучить материал для запоминания? Сколько, например, времени у них уходит на изучение материала, прежде чем они решают, что готовы отвечать, и насколько рационально они распределяют свои усилия по запоминанию? Возможно, одной из причин того, что старшие дети запоминают лучше младших, является более рациональное использование времени, отпущенного на запоминание.

Для проверки своей гипотезы Дюфресн и Кобасигава выбрали методику *парных ассоциаций*. Выполнение задания осуществляется в два этапа: на первом предъявляется ряд из пар слов, а на втором — предъявляется только один элемент каждой пары и ребенок должен попытаться вспомнить недостающий элемент. В своем исследовании Дюфре и Кобасигава использовали два типа пар: «легкие», элементы которых были связаны по смыслу (например, кошка—собака, туфля—носок), и «трудные», элементы которых не имели ничего общего (например, книга—лягушка, коньки—младенец). Всем испытуемым (ученикам 1-х, 3-х, 5-х и 7-х классов) предъявлялись оба набора, и в обоих случаях на предварительное изучение материала давалось неограниченное время.

В табл. 2.1 представлены средние временные показатели при запоминании пар обоих типов. На основании этих данных можно сделать несколько выводов. Как и следовало ожидать, старшие дети обычно дольше изучали материал. Неудивительно и то, что на запоминание «трудных» пар им требовалось больше времени. Наконец, у самых младших испытуемых разница между запоминанием легких и

трудных пар не наблюдалась; только дети из двух старших возрастных групп уделяли больше внимания трудным заданиям. В конце главы мы еще вернемся к последнему на выводу.

Второе исследование также было направлено на изучение памяти, но в других возрастных группах. Черри и Парк (Cherry, Park, 1989) изучали память на пространственную локализацию у выборки молодых (средний возраст 21 год) и пожилых (средний возраст 69 лет) людей. Сначала испытуемым показывали определенную расстановку 24-х обычных предметов, затем предметы перемешивали, а задачей испытуемых было разложить все предметы по прежним местам. Сравнивались результаты при двух типах экспериментальных условий. Половина испытуемых видела предметы на фоне цветной объемной модели со множеством ориентиров. Другая половина — двухмерную черно-белую схему объемной модели. Исследователей интересовало два вопроса. Лучше ли пространственная память в ситуации, когда нужно запомнить местоположение объектов, находящихся в естественных, детализированных условиях? И (как говорят некоторые исследования) будет ли облегчающий эффект удобного контекста у пожилых испытуемых выше, чем у молодых?

В табл. 2.2 представлены полученные результаты. Можно заметить, что контекст действительно имел значение: выполнение на фоне модели было лучше, чем на фоне схемы. Имел значение и возраст: молодые испытуемые превосходили пожилых. Наконец, хотя средняя результативность групп различалась, реакция на изменения контекста была общей. И молодые, и пожилые лучше справлялись с заданием при наличии объемной модели.

Таблица 2.1

Среднее время ознакомления с материалом (в секундах) у детей
из исследования Дюфресна и Кобасигавы
(A. Dufresne & A. Kobasigawa, 1989, *Journal of Experimental Child Psychology*, 47, p. 282)

Тип материала			
Группа	«Легкие» пары	«Трудные» пары	Среднее
1-й класс	5,40	5,20	5,30
3-й класс	5,53	6,96	6,25
5-й класс	4,23	8,42	6,33
7-й класс	4,45	12,48	8,47
Среднее	4,90	8,27	

Таблица 2.2

Среднее количество предметов, правильно расставленных испытуемыми
из исследования Черри и Парк
(K. E. Cherry & D. C. Park, 1989 *Psychology and Aging*, 8, p. 520)

Контекст			
Группа	Схема	Объемная модель	Среднее
Молодые	14,5	15,9	15,2
Пожилые	11,1	14,5	12,8
Среднее	12,8	15,2	

Переменные

Начнем рассмотрение общих принципов с введения некоторых терминов. Психологическое исследование связано с переменными и отношениями между ними. Существует два вида переменных: зависимые и независимые. **Зависимые переменные** — результирующие — те, показатели которых составляет результат исследования. В первом примере зависимой переменной было количество секунд, в течение которых ребенок заучивал каждую пару; во втором примере зависимой переменной было количество предметов, положение которых смог восстановить по памяти каждый испытуемый. Эти переменные являются зависимыми в том смысле, что вариации их значений зависят от других факторов. Центральной задачей исследователя является выяснение этих факторов. Они в любом случае изменяемы: если бы колебание значений зависимой переменной было бы невозможно, не было бы смысла и проводить исследование.

Зависимую переменную исследователь измеряет, но не контролирует непосредственно. **Независимые переменные**, напротив, находятся под контролем исследователя. Цель исследования — определить, действительно ли изменения значений выбранных независимых переменных обуславливают изменения в значении зависимой переменной. В исследовании, которое провели Дюфресн и Кобасигава, независимыми переменными были возраст ребенка и контраст трудного и легкого материала для запоминания, тогда как в исследовании Черри и Парк таковыми были возраст и тип контекста. Эти переменные независимы в том смысле, что их величина заранее определена, а не является результатом исследования. Опять же «изменчивость» необходима: если независимая переменная неизменна, нет никакой возможности выяснить, оказывает ли она воздействие. Изменение и сравнение — неотъемлемая часть любого исследования.

Разделение переменных на зависимые и независимые оправдано во многих, но не во всех исследованиях. Предположим, вы хотите узнать, есть ли связь между *IQ* ребенка и его успеваемостью в школе. Вы могли бы протестировать выборку школьников и получить два показателя: результаты *IQ* теста и школьные отметки. Здесь интересен вопрос, связаны ли изменения одного показателя их с изменениями другого, учатся ли лучше дети с высоким *IQ*? В таком исследовании нет зависимой переменной, величину которой мог бы контролировать экспериментатор; *IQ*, отметки и связь между ними — все это результирующие переменные. Особенности такого «корреляционного» исследования детально рассматриваются далее. Сейчас же важно то, что не для всех исследований подходит схема независимая переменная — зависимая переменная.

Приведенные примеры исследований могут служить в качестве иллюстрации другой особенности независимых переменных. Различные значения независимой переменной можно задать двумя способами. Первый — через экспериментальные манипуляции, которые в буквальном смысле создают переменную. Это то, к чему прибегли Дюфресн и Кобасигава, составляя свои трудные и легкие пары, а также Черри и Парк, создавая плоский и объемный фон. Однако с другой независимой переменной из обоих исследований — хронологическим возрастом — дело обстоит иначе. Очевидно, что исследователи не могут противопоставить два возраста так

же, как легкие и трудные задания. С такими переменными контроль осуществляется не через манипуляции, а через *отбор* испытуемых с желательными характеристиками (например, в возрасте 20 или 70 лет). Поскольку отбор является единственным возможным способом контроля, возраст и другие «субъектные переменные» могут создавать особую проблему при интерпретации — к этому вопросу мы еще вернемся.

Прежде чем идти дальше, введем несколько новых терминов. Независимые переменные также называют **факторами**, а конкретные значения, которые они принимают, — **уровнями**. Поэтому исследование Дюфресна и Кобасигавы можно охарактеризовать как 4×2 -факторное, то есть эксперимент с двумя факторами, один из которых имеет 4 уровня (возраст), а другой — 2 уровня (условия). Аналогично, исследование Черри и Парк можно охарактеризовать как $2 (возраст) \times 2 (условия)$ -факторное. Заметьте, что такого рода обозначение формата говорит нам о количестве экспериментальных ячеек или групп. Например, в исследовании Черри и Парк было $4 (2 \times 2)$ отдельные группы: молодые люди, видевшие модель, молодые люди, видевшие схему, пожилые люди, видевшие модель, и пожилые люди, видевшие схему.

Валидность

В каждом исследовании есть переменные и отношения между ними. Поэтому, когда мы хотим описать исследование, центральным является конструкт переменных: какого вида различия изучаются и какими способами? Если мы хотим выйти за рамки описания и перейти к *оценке* исследования, центральным конструктом становится **валидность**. Вопрос валидности — это вопрос точности: направлено ли исследование на изучение того, что, как предполагается, изучается с его помощью? Все обсуждаемые в книге частные методологические моменты возвращают нас к базовому вопросу точности выводов, которые делаются на основе исследования.

Существует множество форм валидности (Cook & Campbell, 1979). В этой главе рассматриваются три: *внутренняя, внешняя и конструктная*. В главе 7 к ним добавляется еще одна форма: *валидность статистического вывода*.

Внутренняя валидность характеризует внутренний контекст исследования. Речь идет о том, действительно ли независимые переменные связаны с зависимыми так, как это предполагается. Получим ли мы правильные выводы о причинном воздействии (или отсутствии причинного воздействия) одних переменных на другие? Возьмем в качестве примера исследование Дюфресна и Кобасигавы. Их выводы внутренне валидны, если трудные пары действительно требовали больше времени на запоминание, если среднее время запоминания действительно увеличивалось как функция от возраста и если способность отличать легкие пары от трудных действительно улучшалась с возрастом. Если существует альтернативное объяснение любого из этих выводов, тогда внутренняя валидность исследования ставится под сомнение. Предположим, что испытуемые 7-классники были бы в основном из классов для «одаренных» детей, а младшие дети обладали бы средними способностями. Тогда у нас было бы альтернативное объяснение кажущемуся улучшению с возрастом: различия отражают не естественные возрастные изменения, а разницу

в уровне способностей. (Эта проблема называется «систематической ошибкой отбора», далее она обсуждается подробнее.)

Вопрос **внешней валидности** — это вопрос возможности обобщения. Поэтому он встает, как только мы выходим за рамки конкретного исследования. Теперь важно, правомерны ли полученные выводы для другой выборки, других ситуаций и другого поведения — не для любых, конечно, выборок, ситуаций и поведения, а для тех, результат изучения которых, по нашему мнению, должен быть прогностичным. Возьмем, к примеру, исследование Черри и Парк. Их результаты внешне валидны, если у молодых людей действительно в целом лучше пространственная память, а также, если и молодым и пожилым действительно в целом помогает отчетливость и выразительность фона. В обоих случаях «в целом» означает то, что данное явление обнаруживается во всех выборках молодых и пожилых людей, при любых способах измерения пространственной памяти и при любых признаках фона. Если какой-то из результатов нельзя принять за правило по какому-либо параметру, тогда этот результат внешне невалиден. Возможно, к примеру, что характеристики фона имеют значение только в «мелкомасштабной» среде, таких как масштабы предъявлявшиеся Черри и Парк модели и схемы, а также, что этот эффект не сравним с ситуацией естественных размеров предметов обстановки в реальной жизни. Если бы это ограничение имело место (в другом исследовании было показано, что это не так), тогда исследование Черри и Парк ограничивало бы внешнюю валидность.

Добротное исследование должно отвечать требованиям и внутренней, и внешней валидности. Как замечают Кэмпбелл и Стэнли (Campbell & Stanley, 1966), «внутренняя валидность — это базовый минимум, без которого невозможно интерпретировать ни один эксперимент» (р. 5). С точки зрения логики, при отсутствии внутренней валидности невозможно обобщить результаты исследования, поскольку они вообще невалидны. Однако внешняя валидность также играет огромную роль. Внутренне валидные выводы бессмысленны, если они не работают за пределами экспериментальной ситуации.

Внутренняя валидность также является необходимым условием третьей формы валидности — **конструктивной**. Конструктивная валидность имеет отношение к теоретической точности: правильно ли мы объясняем причинно-следственные связи, которые были обнаружены в исследовании? Иными словами, если допустить, что выводы внутренне валидны, знаем ли мы, *почему* получили определенные результаты? Предположим, мы уверены в том, что манипуляции с контекстом в исследовании Черри и Парк действительно обуславливают различия в успешности запоминания. Почему контекст имел значение? Вероятно, наиболее очевидное объяснение — которым руководствуются в большинстве подобных исследований — заключается в том, что важны именно отчетливость и детальность визуальной информации: местоположение предмета лучше всего запоминается тогда, когда он помещен в достаточно дифференцированную среду. Однако вероятно и другое обоснование. Возможно, модель была более интересна и привлекательна, чем схема, что вызывало более пристальное внимание, а отсюда лучшее запоминание. Согласно этой точке зрения, любые манипуляции, которые повышают уровень внимания, должны повышать и успешность выполнения задания вне зависимости от

отчетливости фона или наличия ориентиров. Возможно также, что испытуемые чувствовали себя более уверенно перед пространственной моделью, так как она более привычна, чем абстрактная схема, и именно повышение уверенности улучшало результаты. Если правдоподобные альтернативные объяснения нельзя исключить, тогда исследование характеризует недостаточная конструктивная валидность.

Сказанное выше — всего лишь предварительные замечания о конструктах, которые в дальнейшем мы будем рассматривать с разных точек зрения. Пока же сделаем еще одно замечание, касающееся валидности. Оно имеет отношение к трудности достижения в одном исследовании всех форм валидности. Эта трудность обусловлена тем, что меры, принимаемые исследователем с целью повышения одной формы валидности, работают против другой формы. Необходимость компромиссного выбора наиболее очевидна в отношении внутренней и внешней валидности. В целом, чем жестче контроль, тем выше внутренняя валидность исследования — то есть тем больше уверенности может быть у экспериментатора в том, что переменные связаны между собой именно так, как он предполагает. В то же время искусственный характер ситуации жестко контролируемого эксперимента ставит под угрозу возможность перенесения результатов на внелабораторные условия. И наоборот, исследование, проведенное в естественной обстановке при естественном поведении, не ставит под сомнение возможность обобщения, поскольку ситуации, на которые исследователь хочет перенести свои выводы, максимально приближены к тем, которые он изучает. Однако недостаток экспериментального контроля может крайне затруднить установление валидных отношений.

Отбор

При принятии решений, касающихся переменных, нужно отвечать на вопрос «какие»: какие независимые переменные я буду изменять и какие потенциальные следствия этих изменений я буду измерять? Но важен также ответ на вопрос «кто»: на ком я буду изучать связи между независимыми и зависимыми переменными?

Формирование группы участников эксперимента называется **отбором**. Отбор важен по причине ограниченности рамок исследования. За очень редкими исключениями психологи не имеют возможности обследовать всех людей, которые их интересуют. Изучающий период младенчества, к примеру, не будет обследовать всех младенцев в мире, даже всех, родившихся в США, или (вероятно) даже тех, кто родился в определенном регионе. Вместо этого он обследует **выборку**, на основе результатов которой он надеется сделать общие выводы об интересующей его **популяции**. Обобщение правомерно, если выборка *репрезентативна* в отношении популяции. Это, как явствует, вопрос внешней валидности.

Как исследователю убедиться в том, что выборка репрезентирует популяцию? Следуя логике, сначала нужно определить, что представляет из себя сама популяция. Она необязательно должна быть столь велика, как все младенцы мира, скорее всего это что-то вроде «всех доношенных 3-месячных младенцев США». Следующим шагом после очерчивания границ популяции является **случайный отбор**. Как подразумевает сам термин, случайный отбор означает, что представители популяции имеют равные шансы попасть в выборку. При этом наиболее вероятно, что

в результате процесса отбора характеристики выборки будут отражать характеристики популяции. Заметьте, однако; что вероятность этого напрямую зависит от размера выборки. Случайная выборка, состоящая из 100 человек, будет значительно более репрезентативна, чем случайная выборка, состоящая из 10 человек. Этот принцип — лишь один из ряда доводов (мы столкнемся с некоторыми из них в главе 7) в пользу больших выборок.

Насколько репрезентативны выборки, на практике используемые в исследованиях? Некоторое представление об этом дает нам обзор журналов *Child Development* и *Developmental Psychology*. В первую очередь, какую информацию мы получаем об этих выборках? Во всех проанализированных статьях нам сообщают количество участников и их возраст; в 91 % случаев указывается соотношение количества представителей обоих полов. В 70 % случаев предоставляются некоторые сведения о социальном статусе испытуемых и в 72 % — об их расе. Сравнение с результатами контент-анализа в первом издании книги (которые основывались на анализе тех же журналов за 1983 год) говорит о том, что в последние годы исследования отчетов стали информативнее. Соответствующими показателями из обзора 1983 года были 75 % для пола, 52 % для социального статуса и 24 % для расы. Тем не менее неопределенность при описании расового или социально-экономического состава выборки является скорее правилом, чем исключением. Например, типичны фразы «преимущественно белые», «преимущественно среднего класса»; реже можно встретить точные численные значения и способы подсчета. Опубликованные в недавнем номере *Child Development* критерии рас и социальных классов (Entwisle & Astone, 1994) могут со временем привести к конкретизации исследовательских отчетов.

Неопределенность характерна и для сведений о способе привлечения испытуемых. Как отмечалось в главе 1, испытуемые не появляются по волшебству, исследователь постоянно должен задаваться вопросом, к кому и как обратиться. Имеют значение три момента. Во-первых, каково было исходное количество кандидатов, к которым обращались с просьбой принять участие в эксперименте? Изучая период младенчества, к примеру, использовал ли исследователь объявления о рождении детей в местной газете, списки в роддомах или информацию, полученную у педиатров? Популяции, которые можно выделить, пользуясь этими источниками, могут существенно различаться. Второй вопрос касается процента тех, кто в итоге согласился на участие. Если исследователь обратился, к примеру, к 100 новоиспеченным матерям, сколько из них принесли своих малышей на обследование? И наконец, каков процент испытуемых, прошедших все исследование, — или, иначе говоря, сколько испытуемых отказалось сотрудничать в ходе исследования? В некоторых видах исследования, практически во всех, где фигурируют младенцы, выбывание испытуемых может явиться довольно серьезной проблемой. И тем не менее информация об отборе испытуемых и количестве выбывших встречается в отчетах редко. Только в 34 % статей сообщаются более или менее точные сведения о способе отбора и о доле согласившихся из числа тех, к кому обратились с просьбой об участии. Только в 50 % соответствующих случаев дается информация о том, сколько участников выбыло в ходе исследования.

Учитывая эти недостатки в отчетах, что можно сказать о выборках при изучении развития? Эти выборки, как правило, достаточных размеров, обычно включают лиц и мужского, и женского пола. Испытуемые преимущественно из среднего класса, жители городов, в которых располагаются университеты, а значит, и исследовательские лаборатории. Наконец, в выборки в большинстве своем входит белое население, за исключением тех случаев, когда исследование направлено на выявление возможных групповых различий между белыми и афро-американскими детьми. В этом отношении выводы из этого обзора повторяют выводы контент-анализа других психологических исследований, проведенных за последнее время (Graham, 1992; Hagen & Conley, 1994)¹.

О чем говорит этот обзор? Очевидно, что изучаемые психологами выборки зачастую по тем или иным признакам не соответствуют критерию репрезентативности. Однако насколько серьезны последствия этих несоответствий? Действительно ли они ставят под сомнение внешнюю валидность большей части данных, накопленных исследователями в области психологии развития? По многим причинам ответ, скорее всего, отрицательный. Поскольку параметры практически любой выборки так или иначе будут отклоняться от абсолютной репрезентативности, вопрос может быть сформулирован так: существует ли вообще вероятность того, что определенные отклонения отражаются на возможности обобщения результатов? Часто такая вероятность минимальна. Обобщение предполагает нечто большее, чем то, о чем с логической точки зрения свидетельствуют полученные данные, и во многих случаях гораздо разумнее обобщить, чем не обобщать. Тем не менее к репрезентативности не стоит относиться излишне легкомысленно. «Часто» не то же самое, что «всегда», а «вероятный» не то же самое, что «гарантированный». Характер выборки всегда является тем, что важно учитывать, планируете ли вы свой эксперимент или оцениваете чужой.

Контроль

В предыдущих разделах мы уже затрагивали понятие контроля. Вспомним, что независимая переменная определяется как переменная, которую контролирует исследователь. А обеспечение репрезентативности выборки является одним из видов контроля, который должен осуществлять исследователь. Целью этого раздела является рассмотрение остальных форм контроля, которые приобретают значение с момента получения выборки.

Как явствует из табл. 2.3, для осуществления исследования важны три формы контроля. В таблице перечисляются эти формы и приводятся примеры их использования в исследовании. Формы и примеры их использования довольно сложны; их суть прояснится по ходу изложения. Таблица предназначена для общей ориентации, благодаря которой можно увидеть вводимые разграничения.

Один из видов контроля имеет отношение именно к конкретной форме переменной. Если, к примеру, интерес представляет эффект определенного подкрепления, тогда исследователь должен иметь возможность применить именно такое под-

¹ Я должен добавить, что в последние годы *Child Development* выпускает специальные номера: один посвящен детям из этнических меньшинств (McLoyd & Spencer, 1990), другой — детям из малообеспеченных семей (Huston, McLoyd, & Coll, 1994). Эти специальные выпуски призваны стимулировать изучение недостаточно исследованных тем.

крепление. Если имеют место какие-либо непредусмотренные отклонения — в форме, хронометраже, последовательности и т. д., — исследователь больше не сможет с уверенностью сказать, что представляет из себя независимая переменная. Или рассмотрим вновь параметр времени изучения в исследовании Дюфресна и Кобасигавы. Поскольку исследователей интересовал возможный эффект трудности запоминаемого материала, необходимым условием являлось предъявление всем детям одних и тех же пар, отличавшихся именно по трудности. Вывод, который отсюда можно сделать, незамысловат; он заключается в том, что, желая исследовать эффект чего-либо, следует прежде всего получить возможность сгенерировать это «что-либо». Однако заметьте, что сделать это не всегда так просто, как в приведенных примерах, в которых уровни независимой переменной определялись всего лишь различием стимульного материала. Если экспериментальные манипуляции сложнее, использование переменной в одном и том же виде со всеми испытуемыми может оказаться проблематичным.

Вторая форма контроля имеет отношение ко всем факторам экспериментальных условий, за исключением независимой переменной. Независимые переменные не находятся в вакууме; для них всегда существует контекст, и задачей исследователя является точное определение того, что это будет за контекст. Проверая память, (как, например, в двух описанных выше исследованиях), исследователь должен решить не только то, какой использовать тест, но и то, какова будет обстановка тестирования. Простой выход в этом конкретном случае — сделать обстановку как можно более спокойной с тем, чтобы минимизировать отвлекающие моменты. Приняв это решение, экспериментатор берет на себя ответственность по обеспечению равно спокойной обстановки для каждого испытуемого.

Теперь введем еще несколько терминов. Вариации значения зависимой переменной называются *дисперсией*. Те вариации, которые можно объяснить действием независимой переменной, называются *первичной дисперсией*; те же, которые являются результатом действия других факторов, называются *вторичной дисперсией*, или *дисперсией ошибки*. Контролируя уровень других потенциальных переменных, экспериментатор пытается максимизировать долю первичной дисперсии. Возможно даже большее значение имеет то, что он стремится свести к минимуму вероятность устойчивой связи между любой из независимых переменных и другими источниками дисперсии. Допустим, что Черри и Парк тестировали бы всех своих молодых испытуемых в тихой университетской лаборатории, а пожилых испытуемых — в шумной комнате клуба пенсионеров. Очевидно, что в этом случае были бы две независимые переменные — возраст и условия тестирования, — тогда как подразумевалось наличие только одной. Любые непредусмотренные связи между двумя потенциально значимыми переменными называются *смешением*. Одной из первоочередных целей качественного исследовательского плана является исключение возможности смешения.

Как явствует из табл. 2.3, контроль нежелательных переменных может принимать несколько форм. Нередко имеется возможность контролировать переменную, делая ее одинаковой для всех испытуемых. Этот случай иллюстрирует пример с памятью, в котором шум в экспериментальных условиях поддерживался на одном уровне для всех испытуемых. Однако иногда такое точное уравнивание нецелесообразно. За примером можно вновь обратиться к исследованию Дюфресна и Кобасигавы. На успешность выполнения теста школьниками могло влиять время суток,

в которое проходило тестирование. Склонность к сотрудничеству и внимание в конце учебного дня не обязательно те же, что и перед уроками, непосредственно перед переменной — не те же, что сразу после нее, в пятницу — не те же, что в понедельник. Очевидно, что Дюфресн и Кобасигава внесли бы потенциально значимое смещение, если бы тестировали всех своих первоклассников утром, а всех 7-классников днем. Один из способов избегания этой проблемы — тестирование всех детей в одно и то же время суток в один и тот же день недели, к примеру, в час дня в среду. Однако с таким подходом для завершения большинства исследований потребовались бы месяцы, и даже тогда постоянными для всех испытуемых оставались бы лишь время суток и день недели, а не время года (которое тоже может иметь значение). Разумным альтернативным ходом было бы допустить колебания времени тестирования, однако сделать так, чтобы эти колебания были бы одинаковы во всех сравниваемых группах — в данном случае 1-, 3-, 5- и 7-классников. В этой ситуации контроль переменной «время тестирования» состоял бы не в приравнивании, а в рандомизации — то есть в равномерном распределении ее вариаций по группам испытуемых.

Таблица 2.3

Формы экспериментального контроля

Тип контроля	Методы достижения	Примеры
Независимой переменной	Сделать значимые элементы экспериментальных манипуляций одинаковыми для всех испытуемых	В эксперименте Дюфресна и Кобасигавы одним и тем же способом предлагать всем детям один и тот же набор легких и трудных пар
Потенциально значимых факторов экспериментальных условий	Сохранять факторы постоянными для всех испытуемых	В эксперименте Черри и Парк тестировать всех испытуемых в одной тихой комнате
	Равномерно распределить по группам вариации других факторов	В эксперименте Дюфресна и Кобасигавы случайным образом выбирать время тестирования во всех группах детей
Изначальных индивидуальных различий между испытуемыми	Случайным образом причислять испытуемых к группам, обследуемым в разных экспериментальных условиях	В эксперименте Черри и Парк случайным образом включить половину испытуемых обоих возрастов в группу наблюдающих модель и в группу наблюдающих схему
	Подобрать испытуемых таким образом, чтобы они соответствовали друг другу по потенциально значимым качествам	В эксперименте Черри и Парк измерить IQ испытуемых и причислить испытуемых с равными IQ к разным группам (на практике этого сделано не было)
	Провести обследование каждого испытуемого при всех экспериментальных условиях	В эксперименте Дюфресна и Кобасигавы протестировать каждого ребенка с использованием и трудного, и легкого экспериментального материала

Общие принципы, о которых говорилось выше, должны казаться знакомыми. То, о чем здесь говорится, это просто классический научный метод: выявить эффекты некоторого фактора, систематически изменять этот фактор (первая форма контроля), сохраняя постоянными другие потенциально значимые факторы (вторая форма контроля).

Существует и третья форма контроля, которая также играет важную роль. До сих пор «другие, потенциально значимые факторы», о которых шла речь, находились внутри экспериментальной ситуации, например уровень шума в комнате для тестирования. В любом эксперименте еще одним значимым источником дисперсии являются индивидуальные различия между испытуемыми. Испытуемые заведомо неодинаковы, и различия между ними обуславливают наличие дисперсии ошибки в конечных результатах. Поскольку этих различий нельзя избежать, метод контроля вновь должен заключаться не в уравнивании, а в распределении. Экспериментатор должен убедиться в том, что различия равномерно распределены по группам — или, говоря это же иными словами, что к моменту начала обследования группы эквивалентны. Для выполнения этого требования необходимо, чтобы экспериментатор контролировал не только экспериментальное воздействие, но и тех, на кого оно направлено.

Как экспериментатору распределить людей по группам таким образом, чтобы эти группы оказались изначально эквивалентными? Ответ состоит в том, что, хотя возможность стопроцентной гарантии эквивалентности отсутствует, есть определенные способы приблизиться к ней настолько, насколько разумно было бы ожидать. Наиболее распространенный метод — **случайное причисление** испытуемых к разным группам. Случайное причисление означает, что у всех испытуемых равные шансы попасть в каждую группу. При этом характеристики каждого испытуемого (*IQ*, пол, опыт выполнения подобного задания — все, что может отразиться на результатах) с одинаковой вероятностью могут попасть в любую из групп. Отсюда, результатом случайного причисления, скорее всего, будет равномерное распределение характеристик по группам, что, естественно, и является целью исследования. Ясно, что логика случайного причисления та же, что и логика случайного отбора, и успешность этого процесса также зависит от размера выборки. Нельзя произвольно разделить 8 испытуемых на две группы и считать с какой-либо долей уверенности, что в результате рандомизации были получены эквивалентные группы. С выборкой, состоящей из 80 человек, шансы на успех гораздо выше.

На практике исследователи в области психологии развития, формируя группы, редко используют абсолютно случайное причисление. Обычно исследователь стремится работать с уравненными экспериментальными группами. Независимо от размеров выборки случайное причисление не может гарантировать, что к концу исследования количество испытуемых не уменьшится. Однако можно поставить условие, согласно которому размеры группы к концу исследования должны быть равными, при этом сохраняя случайность причисления всех испытуемых. Аналогично, даже если исследователя не интересуют половые различия, имеет смысл обеспечить равное соотношение мальчиков и девочек во всех сравниваемых группах. Или, проводя обследование в школе, полезно удостовериться в равноценности обстановки в разных классных комнатах. Соответствие любому из этих требо-

ваний можно гарантировать установлением определенных ограничений на случайное причисление.

При обсуждении необходимости ограничения влияния случайности встает закономерный вопрос: зачем вообще использовать случайное причисление? В конечном счете цель исследователя — обеспечить изначальную эквивалентность групп, но в то же время мы видим, что случайность не гарантирует эквивалентность. Если мы можем уравнивать соотношение полов испытуемых и обстановку классных комнат, почему не пойти дальше и не привести в соответствие *все* имеющие значение характеристики, таким образом *обеспечивая* эквивалентность групп? Общий ответ заключается в том, что привести в соответствие все параметры — труднее, чем может показаться на первый взгляд, и попытка сделать это иногда усложняет, а не упрощает дело. Более конкретный ответ дан в главе 3, где мы возвращаемся к вопросу об отборе и причислении испытуемых. Кроме того, в главе 3 рассматривается третий метод достижения эквивалентности: обследование каждого испытуемого во всех экспериментальных условиях.

Субъектные переменные

Переменные, поддающиеся и не поддающиеся манипуляциям

До сих пор при обсуждении вопроса экспериментального контроля мы говорили в основном об идеальной для исследователя ситуации: когда он имеет возможность систематически изменять независимые переменные, сохраняя постоянными другие факторы, и включать испытуемых в группы с разными экспериментальными условиями, либо случайным образом, либо произвольно, но в рамках определенных ограничений. При использовании многих переменных такой контроль не только желателен, но и вполне осуществим. Мы видим элементы подобного рода контроля в обоих описанных исследованиях: контраст «легкое—трудное» в эксперименте Дюфресна и Кобасигавы, а также контраст «модель—схема» в эксперименте Черри и Парк.

Однако жизнь психолога осложняется тем фактом, что не все переменные находятся во власти исследователя, чего требует хороший исследовательский план. И вновь оба описанные исследования иллюстрируют эту ситуацию, и в этом случае примером служит хронологический возраст. Ясно, что возраст — не является характеристикой, устанавливаемой исследователем; наоборот, это характеристика самого человека, влияющая на экспериментальные условия. Возраст — это лишь один пример того, что называется **субъектными** (или **классификационными**) **переменными**: неотъемлемыми качествами испытуемых, не поддающимися экспериментальным манипуляциям; качествами, которые должны учитываться в своем естественном виде. Другими характерными примерами являются расовая и половая (если не учитывать успехи хирургии последних лет) принадлежность. Как уже отмечалось, исследователь, который хочет работать с такими характеристиками в качестве независимых переменных, лишает себя возможности контролировать их через манипуляции. Единственный способ контроля в этих случаях — контроль через отбор испытуемых, уже обладающих нужными характеристиками.

Другие переменные хотя и не являются в буквальном смысле не поддающимися манипуляциям, в экспериментах с людьми фактически никогда не контролируются. С теоретической точки зрения, к примеру, было бы весьма интересно узнать, развиваются ли младенцы без матери так же, как и с матерью. За исключением эксперимента Фридриха II (см. сноску на с. 13), у нас нет исследований, посвященных этому вопросу, с использованием манипуляций. Тем не менее уже давно существует литература по «материнской депривации» и ее воздействию на ребенка. Работа исследователей заключалась в том, чтобы выявить ситуации, в которых младенцы остались без матери (обычно в приютах), а затем воспользоваться этими «естественными условиями» для изучения развития детей. Существует масса примеров, когда психологи «эксплуатировали» естественно сложившиеся обстоятельства — исследования недостаточного питания в младенчестве, отсутствия отца в детском возрасте, социальной изоляции в старости и т. д. Во всех случаях независимая переменная создается через отбор, а не посредством манипуляций.

Исследования с не поддающимися манипуляциям переменными не соответствуют критериям «подлинного эксперимента», поскольку исключают возможность контролируемых манипуляций, составляющих суть эксперимента. По этой причине Кемпбелл и Стэнли (Campbell & Stanley, 1966), рассматривая экспериментальный план, называли такие исследования *доэкспериментальными*. Из-за недостатка контроля мы не можем говорить об установленных причинно-следственных связях с той же долей уверенности, что и в классическом эксперименте.

Каковы конкретные минусы исследования с переменными, не поддающимися манипуляциям? Проблемы можно разделить на две основные категории. Во-первых, испытуемых нельзя произвольно распределять по группам. Поскольку случайное причисление невозможно, нельзя быть уверенными и в том, что изучаемые группы эквивалентны не только по интересующей нас переменной (например, наличие или отсутствие матери), но и по другим параметрам, и поэтому нельзя быть уверенными в том, что причиной различий между группами является именно эта переменная. Данное обстоятельство фактически и являлось поводом для критики исследований ранней материнской депривации. Отбор из популяции только тех малышей, которые воспитываются в приютах, не соответствует критерию случайности, и в результате формируется группа детей, в которой, как правило, выше процент генетических и органических нарушений. Поэтому отличие детей из приюта от остальных нельзя с какой-либо долей уверенности отнести на счет эффекта воспитания без матери. В хорошо спланированном эксперименте такое смещение должно было бы исключаться процедурой случайного причисления. Это, очевидно, проблема с внутренней валидностью: мы не можем утверждать, что наша независимая переменная действительно является причинным фактором.

Другая трудность связана с тем, что воздействие большинства субъектных переменных носит множественный и длительный характер. Воспитание в приюте, отсутствие отца, социальная изоляция, черный (или белый) цвет кожи, принадлежность к мужскому (или женскому) полу — все это факторы, которые способны оказывать значимое влияние на развитие человека. Поэтому, даже обнаружив значимый эффект, связанный с определенной субъектной переменной, мы все еще не

будем знать, каковы конкретные причинные факторы. Это еще одна слабая сторона исследований материнской депривации. Хотя пагубные последствия определенных аспектов воспитания в приюте не вызывают сомнений, уже долгое время ведутся споры о том, являются они результатом нехватки материнского тепла (как, например, утверждал Боулби (Bowlby, 1952)), или более общей когнитивно-перцептивной депривации (как утверждал, например, Кэслер, (Casler, 1961, 1961)). Даже если бы мы могли сделать вывод о значимости матери самой по себе, мы все еще не знали бы, отсутствие какого из факторов, связанных с наличием матери, обуславливает обнаруженные негативные последствия. Вновь имеет место смешение факторов, которые разграничиваются в хорошо спланированном эксперименте. Исследователь, держащий переменные под контролем, вряд ли выберет в качестве независимой такую глобальную переменную, которая не позволит интерпретировать возможные эффекты ее воздействия. Это, очевидно, проблема с конструктивной валидностью: мы не знаем, правильно ли интерпретируем результаты.

Это рассуждение не имело своей целью убедить в том, что демонстрация наличия связи между материнской депривацией, половой принадлежностью или возрастом и развитием ребенка бессмысленно. Однако следует отдавать себе отчет в том, что оно является лишь начальным этапом исследовательской программы.

Возраст в качестве переменной

В связи со своей значимостью в психологии развития такая переменная, как хронологический возраст, заслуживает особого упоминания. Целью многих исследований в области психологии развития является определение того, отличаются или не отличаются испытуемые разных возрастов по изучаемым зависимым переменным. Обзор журналов *Child Development* и *Developmental Psychology* (см. табл. 1.3) свидетельствует о том, что 69 % исследований включали по крайней мере две возрастные группы; в 31 % было три и более возрастных группы. Но эти цифры, естественно, не отражают в полной мере масштабы использования возрастных сравнений, поскольку сравнение между возрастными за частую бывает неявным. Например, исследование новорожденных может не включать сравниваемую группу старших детей, однако результаты будут интерпретироваться исходя из сведений об особенностях старших детей. Если взять простой пример, вряд ли кто-нибудь пытался бы определить, есть ли у младенцев цветное зрение (например, Bornstein, 1978), если бы не был уверен в том, что цветное зрение в конечном счете входит в среду возможностей человека.

Специалисты в области психологии развития иногда оправдываются за то, что большинство исследований направлено «только лишь на выявление возрастных различий». Однако сведения об истинных возрастных изменениях, очевидно, имеют для науки о развитии неоспоримую ценность. Описание — не просто неотъемлемая часть любой науки; точное описание дает в распоряжение исследователя феномен, который должна объяснить теоретическая модель. Только зная о том, например, что маленькие дети не понимают принципа сохранения (Piaget & Szeminska, 1952), мы можем начать выстраивать модель, объясняющую этот факт и его причины.

Хотя нужно согласиться с тем, что изучение возрастных изменений оправдано, важно прояснить, что конкретно подразумевается под «истинными возрастными изменениями». При этом, конечно, *не* подразумевается, что хронологический возраст в буквальном смысле является причиной изменений. *Подразумевается*, что изменения обусловлены влиянием переменных, стабильно и естественно связанных с возрастом. Тогда задача исследователя состоит в определении того, какая из потенциально значимых переменных действительно имеет значение.

Выше мы подчеркивали, что основной целью экспериментального контроля является формирование групп, эквивалентных по всем параметрам кроме исследуемой независимой переменной. Эта цель приобретает особую значимость в случае такой глобальной субъектной переменной, как возраст. Представьте, что вам нужно сравнить 7-летних и 12-летних детей. Если вы хотите добиться эквивалентности групп по всем параметрам кроме возраста, вам придется найти 7- и 12-летних детей одного уровня биологической зрелости, посещавших школу одинаковое количество лет, со сходным жизненным опытом и т. д. Очевидно, что эта цель не только нереалистична, но и неверна. Биологическая зрелость, количество лет обучения в школе и жизненный опыт относят к переменным, «стабильно и естественно связанным с возрастом». В этом качестве их и следует изучать, а не исключать посредством экспериментального контроля.

С другой стороны, есть и другие потенциально значимые факторы, которые могут привести к смещению при сравнении возрастных групп. Произошло бы явное смещение, если бы все 7-летние дети были мальчиками, а 12-летние — девочками. Принадлежность к мужскому полу не является обязательной для 7-летнего ребенка, равно как и принадлежность к женскому полу необязательна для 12-летнего, поэтому следует исключить возможность параллельного изменения этого фактора и возраста. Несколько менее явное смещение произошло бы, если всех 7-летних детей взяли из одной школы, а всех 12-летних — из другой. Факт обучения в разных школах, вероятно, не играет большой роли, и в любом случае это различие может оказаться неизбежным в определенном возрастном диапазоне. Тем не менее важно подбирать школы как можно более близкие по таким параметрам, как подход к обучению, местоположение и социально-экономический статус обслуживаемого населения. Если этот критерий не учитывается, тогда кажущиеся возрастные изменения могут действительно оказаться «ложными».

Как следует из примеров, принять решение о том, что приводить в соответствие при сравнении возрастных групп, а что нет, обычно бывает довольно просто. Однако, как мы увидим, это решение не всегда очевидно, не всегда легко и привести в соответствие то, что намечено. К вопросу о возрастном сравнении мы вернемся в главе 3.

Результаты

Исследователи измеряют независимые переменные с тем, чтобы определить, как при этом изменяется значение зависимой переменной. Но как оно может изменяться? В факторном исследовании — то есть в исследовании с двумя и более независимыми переменными — воздействие независимой переменной может принимать две формы: главный эффект и взаимодействие.

Главный эффект

Главный эффект — это результат непосредственного воздействия независимой переменной. Это то, что интересует исследователей при сравнении показателей по уровням отдельной независимой переменной — независимой от других независимых переменных (или итоговой для всех их значений). Оба исследования, приведенные в качестве примера, иллюстрируют понятие главного эффекта. В эксперименте Черри и Парк главным был эффект возраста: молодые испытуемые справлялись с заданием лучше пожилых. Средние значения для этого эффекта представлены в правом столбце табл. 2.2; это итоговые показатели для всех молодых и всех пожилых испытуемых при разных уровнях другой независимой переменной (модель/схема). Подобно тому имел место и главный эффект экспериментальных условий; значения для этого эффекта представлены на нижней строчке таблицы: это итоговые показатели всех испытуемых в ситуации с моделью и в ситуации со схемой для двух уровней возраста.

В эксперименте Дюфресна и Кобасигавы также был главный эффект возраста и главный эффект экспериментальных условий. Итоговые показатели для этих эффектов представлены в табл. 2.1 в столбце и строчке «среднее». Таким образом, в обоих случаях можно говорить о том, что эффект имели обе переменные: значение зависимой переменной изменялось как функция от возраста и экспериментальных условий. Заметьте, однако, что в исследовании Дюфресна и Кобасигавы эффект возраста сложнее других главных эффектов, поскольку независимая переменная имеет 4 уровня а не 2. Главный эффект переменных, имеющих более двух уровней, представляет особые трудности для статистической обработки и интерпретации — к этой проблеме мы вернемся в главе 7.

Взаимодействие

Главный эффект — это эффект отдельно взятой независимой переменной. **Взаимодействие** же возможно при одновременном учете двух и более независимых переменных. Взаимодействие имеет место всегда, когда эффект одной независимой переменной зависит от уровня другой независимой переменной.

Помимо двух главных эффектов в эксперименте Дюфресна и Кобасигавы есть и взаимодействие. Здесь эффект трудности задания изменяется вместе с возрастным уровнем — незначительный эффект в двух младших и значительный эффект в двух старших группах. Любое взаимодействие можно описать двумя способами: эффект возраста изменялся с уровнем трудности задания — никакого различия при выполнении легких заданий, существенное различие при выполнении трудных заданий. Такое двунаправленное («двунаправленное», поскольку имеются две независимые переменные) взаимодействие в виде графиков изображено на рис. 2.1. Данные те же, что и в табл. 2.1, однако графическое изображение демонстрирует характер взаимодействия более наглядно. Обратите особое внимание на то, что линии непараллельны. Графическим признаком взаимодействия является отклонение от параллельности — расхождение или пересечение графиков, отражающее зависимость эффекта одной переменной от значения уровня другой.

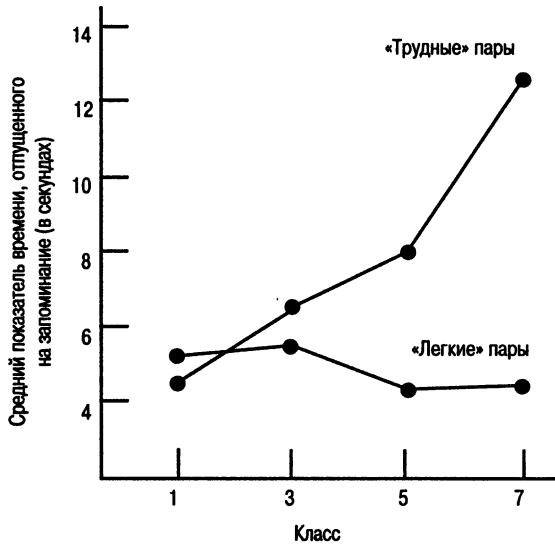


Рис. 2.1. Взаимодействие возраста и экспериментальных условий в исследовании Дюфресна и Кобасигавы.
(A. Dufresne & A. Kobasigawa, 1989, *Journal of Experimental Child Psychology*, 47, 274–296)

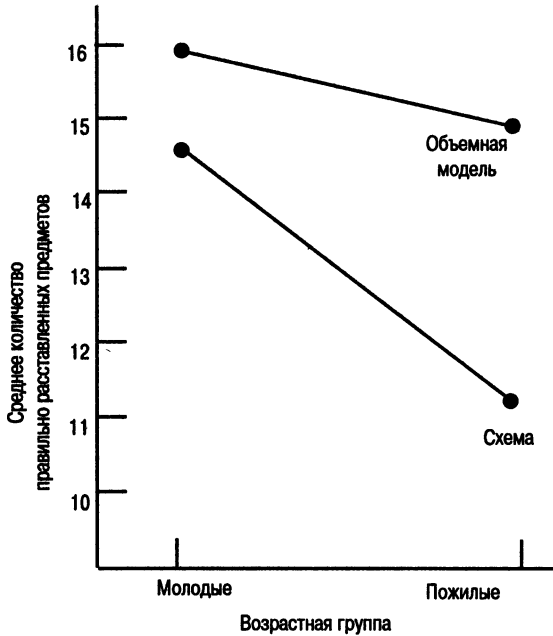


Рис. 2.2. Главные эффекты в исследовании Черри и Парк. (K. E. Cherry & D. C. Park, 1993, *Psychology and Aging*, 8, 515–526)

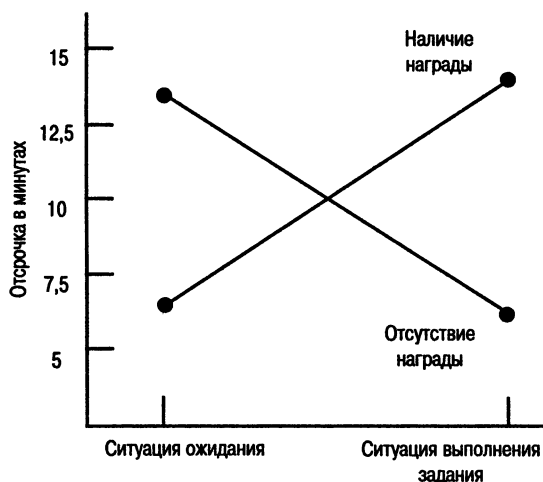


Рис. 2.3. Взаимодействие экспериментальных условий в исследовании Паттерсон и Картер. (C. J. Patterson & D. B. Carter, 1979, *Child Development*, 50, 272–275)

Как бы выглядели графики при отсутствии взаимодействия? Ответом может служить рис. 2.2, на котором представлены графики средних показателей из эксперимента Черри и Парк. Вспомним, что в их исследовании было обнаружено одинаковое улучшение в ситуации с моделью и у молодых, и у пожилых испытуемых — поэтому здесь два главных эффекта (возраст и условия), но нет взаимодействия. Эта ситуация отражена в том, что линии на рис. 2.2 почти параллельны (то, что они не строго параллельны, обусловлено незначительной тенденцией к взаимодействию — тем, что у пожилых испытуемых улучшение в ситуации с моделью более заметно).

В исследовании Дюфресна и Кобасигавы субъектная переменная взаимодействовала с экспериментально манипулируемой. Однако взаимодействие происходит не только по такой схеме; оно может существовать между любыми независимыми переменными. Поэтому его вероятность велика в любом многофакторном эксперименте. Рис. 2.3 иллюстрирует взаимодействие двух экспериментально манипулируемых переменных, а рис. 2.4 — двух субъектных переменных. Главным результатом исследования Паттерсона и Картера (Patterson & Carter, 1979), которое иллюстрируется рис. 2.3, было то, что наличие желаемого вознаграждения ослабляло самоконтроль детей, когда они лишь ожидали его, но усиливало самоконтроль, когда они выполняли ради его получения определенное задание. Одним из результатов исследования, проведенного Андервудом, Койе и Хербсменом (Underwood, Coie & Herbsman, 1992), (см рис. 2.4) явилось выявление изменения склонности детей к демонстративному поведению с целью скрытия грусти, связанного с возрастом и полом испытуемых. В двух младших классах девочки несколько чаще мальчиков говорили о том, что предпочтут скрывать свою грусть; однако в 7-м классе по этой характеристике уже мальчики превосходили девочек.

Интерпретация любого взаимодействия может вызвать трудности и статистические, и теоретические (Levin, 1985; Rosnow & Rosenthal, 1995). Ограничимся

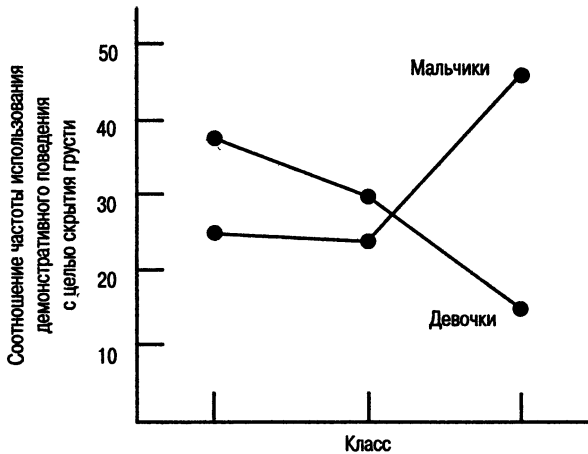


Рис. 2.4. Взаимодействие возраста и пола в исследовании Андервуда, Койе и Хербсмана. (М. К. Underwood, J. D. Coie, & C. R. Herbsman, 1992, *Child Development*, 63, 366–380)

одним важным замечанием. Наиболее общим правилом, которого следует придерживаться в ситуации значимого взаимодействия переменных, является осторожность в интерпретациях. В исследовании Дюфресна и Кобасигавы, к примеру, присутствовал главный эффект и возраста, и трудности задания, однако, как следует из рис. 2.1, эффект возраста проявлялся только при выполнении трудных заданий, а эффект трудности проявлялся только у детей из старших возрастных групп. В исследовании Паттерсона и Картера, напротив, главный эффект наличия/отсутствия вознаграждения был *несущественным*, что свидетельствовало бы о том, что эта переменная не играет роли, если бы этому не противоречили результаты анализа, отдельно — ситуации выполнения заданий и отдельно — ситуации простого ожидания. Таким образом, взаимодействие говорит нам о том, что мир сложнее, чем можно было ожидать. Изучение отдельной независимой переменной не дает полной картины действия переменных.

Причины искажения валидности

Как мы увидели, конечной целью планирования исследования всегда является достижение валидных выводов об изучаемом феномене. Неудачный исследовательский план ставит валидность под сомнение, не устраняя спорных моментов и ограничивая возможности обобщения. В этой главе мы уже коснулись некоторых причин искажения валидности, а в дальнейшем рассмотрим значительно больше. Для ознакомления с последующим текстом может быть полезной краткая сводка факторов, которые следует иметь в виду — общий перечень и ряд определений, к которым можно обращаться при необходимости. Эту роль выполняет табл. 2.4.

Таблица 2.4 составлена по работам Кемпбелла и Стэнли (Campbell & Stanley, 1966), а также Кука и Кемпбелла (Cook & Campbell, 1979). Она не содержит полный перечень помех для исследования (Кук и Кемпбелл описывают 33 вида при-

чин искажения валидности!), однако включает в себя многие из проблем, которые будут обсуждаться по ходу текста. Таблица вряд ли не потребует разъяснений; ее цель — введение понятий, на которые в дальнейшем будет обращено более пристальное внимание.

Таблица 2.4

Причины искажения валидности

Систематическая ошибка при отборе	Отбор для сравнения изначально неэквивалентных испытуемых в группы
Избирательные выбывания из исследования	Неслучайная, систематическое выбывание испытуемых в ходе исследования
История	Потенциально значимые помимо изучаемых независимых переменных события в период между проводимыми измерениями
Созревание	Естественное развитие испытуемых как функция от времени, продолжения исследования
Тестирование	Эффект повторного выполнения одного и того же теста
Реактивность	Непредусмотренное влияние экспериментальной обстановки на ответы испытуемых
Инструментарий	Непредусмотренная смена экспериментаторов, наблюдателей или способов измерения в ходе исследования
Статистическая регрессия	Тенденция приближения к среднему значению первоначально крайних показателей при повторных тестированиях
Низкая надежность	Ошибки измерения при оценке зависимой переменной
Ограниченность статистических возможностей	Низкая вероятность выявления истинных эффектов, обусловленная особенностями экспериментального плана и статистических критериев
Моно-операционное искажение	Использование только одного способа операционализации либо независимой, либо зависимой переменной
Моно-методическое искажение	Использование только одного экспериментального метода для изучения связей между независимыми и зависимыми переменными

Резюме

Глава начинается с введения некоторых базовых терминов и понятий. Все исследования предполагают наличие переменных. *Зависимые переменные* — результаты, например, количество агрессивных действий при изучении агрессии. *Независимые переменные* — это потенциальные причинные факторы, которые контролирует исследователь — например, подкрепление агрессии. Цель большинства исследований — определить, связаны ли изменения независимой переменной с изменением значения зависимой переменной — например, повышается ли уровень агрессии после получения подкрепления?

Главным вопросом во всех исследованиях является вопрос валидности. *Валидность* — это обоснованность выводов, которые можно сделать из данного исследования. В главе рассматриваются три вида валидности: *внутренняя валидность*, которая имеет отношение к точности выводов о причинно-следственных связях внутри контекста исследования; *внешняя валидность*, которая касается возможности генерализовать сделанные выводы; и *конструктивная валидность*, которая имеет отношение к безупречности теоретических интерпретаций и выводов.

Важнейшее решение, которое должен принять исследователь, касается участников исследования. Целью отбора испытуемых является получение *выборки*, репрезентативной в отношении *популяции*, на которую исследователь хочет перенести свои заключения. Общим принципом достижения репрезентативности является *случайный отбор* из интересующей популяции. На практике в большинстве исследований в области психологии развития используется не абсолютно случайный отбор, а большинство выборок по тем или иным признакам не соответствуют критерию репрезентативности. Насколько значимы эти несоответствия, несомненно зависит от изучаемых проблем. Тем не менее репрезентативность и внешняя валидность остаются существенными моментами, которые нужно учитывать в любом исследовании.

Затем было рассмотрено понятие контроля. Под контролем находится независимая переменная. Для получения четкой информации о причинно-следственных связях важно три вида контроля. Первый — контроль точности независимой переменной. Второй — контроль других потенциально значимых факторов в экспериментальной ситуации. Анализируется два метода осуществления второй формы контроля: поддержание на одном уровне других факторов или случайное распределение их значений между испытуемыми. Третий вид контроля — контроль изначальных индивидуальных различий между испытуемыми. Один из методов его осуществления — случайное причисление — рассматривается в этой главе; два других вида (приведение в соответствие и внутрисубъектное тестирование) будут рассмотрены в следующих главах.

В некоторых исследованиях степень контроля ограничена характером переменных. Термин *субъектная переменная* имеет отношение к изначальным различиям между людьми, характеристикам, которыми нельзя манипулировать; в качестве примера можно привести пол, возраст, расу. Единственной формой контроля таких переменных является отбор, что имеет значение также в тех ситуациях, когда экспериментальное воздействие было бы неэтичным (например, материнская депривация). Хотя у специалистов в области психологии развития не поддающиеся манипуляциям переменные зачастую вызывают особый интерес, установление причинно-следственных связей при отсутствии экспериментальных манипуляций довольно затруднительно. Выявление причин наблюдаемых феноменов проблематично, если влияние переменной множественно и длительно; кроме того, бывает трудно исключить возможность наличия других причинных факторов.

Субъектные переменные представляют особый интерес, когда вступают во взаимодействие. *Взаимодействие* определяется как зависимость эффекта одной независимой переменной от уровня другой. *Главный эффект*, напротив, является результатом воздействия одной независимой переменной, не связанной с другими

факторами. Взаимодействие может происходить между независимыми переменными любого вида и принимать разные формы. Взаимодействие говорит о сложности отношений между переменными и о необходимости быть осторожными в своих выводах о любой из них.

В конце главы автор возвращается к понятию валидности. Имеется множества факторов, искажающих валидность, здесь перечислены важнейшие из них. Далее они будут рассматриваться более подробно.

Упражнения

Найдите в популярных изданиях (газетах, журналах) по крайней мере три заметки, знакомящие читателей с результатами каких-либо исследований в области психологии развития. Составьте список возможных причин искажения валидности. Если предоставленной информации недостаточно, чтобы можно было оценить некоторые формы валидности, определите, какие дополнительные сведения вам необходимы.

Подумайте над задачей привлечения испытуемых из следующих возрастных групп: 6 месяцев, 4 года, 12 лет, 70 лет. Для каждой группы составьте перечень способов формирования выборки. Оцените вероятную репрезентативность выборок при каждом методе отбора.

Один и тот же конструкт может выступать в роли и независимой и зависимой переменной, что определяется способом его использования. Возьмем в качестве примера следующие конструкты: тревожность, уровень активности, готовность к школьному обучению. Для каждого из конструктов спланируйте исследование, в котором он выступал бы в роли: а) зависимой переменной; б) независимой переменной; в) субъектной переменной; г) корреляционной переменной.

Предположим, в исследовании две независимые переменные, A и B , у каждой из которых два уровня — то есть план 2×2 . Зависимая переменная C может принимать значение от 0 до 50. Начертите графики, иллюстрирующие каждый из приведенных ниже возможных результатов, а затем объясните, что означают эти результаты:

- а) значимые главные эффекты A и B при отсутствии взаимодействия;
- б) значимое взаимодействие между A и B при отсутствии главных эффектов;
- в) значимый главный эффект A и значимое взаимодействие между A и B .

В главе 2 говорилось о том, что в любом исследовании присутствует сравнение. В большинстве случаев сравниваются показатели при разных значениях независимой переменной. Если независимая переменная — не подающаяся манипуляциям субъектная характеристика, например возраст, тогда исследователь должен подобрать испытуемых, обладающих разными уровнями данной характеристики. Если независимая переменная — это экспериментально манипулируемый фактор, тогда исследователю нужно поставить испытуемых в условия, которые характеризуют определенные значения изучаемого фактора. В любом случае, исследователь должен отбирать испытуемых и включать их в группы способом, позволяющим осуществить объективное сравнение разных уровней (внутренняя валидность), которое дает возможность перенесения выводов на другие интересующие исследователя выборки (внешняя валидность) и выявить причинные основания обнаруженных связей (конструктивная валидность).

Обрисованные здесь этапы и цели реализуются в *экспериментальном плане*. План, по словам Керлингера (Kerlinger, 1986, р. 279), — это «замысел и структура исследования», то, как объединены разные исследовательские приемы. Хотя конечная цель — получение валидных выводов — всегда сохраняется неизменной, на практике отдельные исследовательские приемы можно объединить разными способами. В этой главе рассмотрены некоторые из наиболее значимых параметров, по которым различаются исследовательские планы.

И вновь исследования, приведенные в главе 2 в качестве примера, послужат иллюстрацией некоторых общих принципов и стандартной терминологии. Эксперимент Дюфресна и Кобасигавы, а также эксперимент Черри и Парка включают два уровня манипулируемой переменной: легкий/трудный материал для запоминания у первых и контекст модели/схемы у последних. Черри и Парк ставили каждого из испытуемых в одно из двух экспериментальных условий, поэтому их подход можно назвать **межсубъектным планом**. Дюфресн и Кобасигава тестировали всех испытуемых и в легких, и в трудных условиях; поэтому их подход можно назвать **внутрисубъектным планом**. Одно из кардинальных решений, которое должен принять исследователь, — использовать при сравнении эффектов двух или более эк-

спериментальных воздействий одних и тех же или разных испытуемых. Далее в этой главе будут проанализированы плюсы и минусы обоих подходов.

Оба описанных исследования также включали в качестве неманипулированной переменной хронологический возраст. В этой ситуации был использован общий методический прием: обе группы исследователей тестировали испытуемых разного возраста. Стратегия тестирования разных групп людей разного возраста называется **планом поперечных срезов**. Это не единственный подход к изучению возрастных изменений. Дюфресн и Кобасигава могли бы, к примеру, протестировать выборку первоклассников, подождать два года и протестировать тех же детей уже как третьеклассников, подождать еще два года и протестировать их уже как пятиклассников и, наконец, еще через два года — как семиклассников. Стратегия регулярного повторного тестирования одной выборки испытуемых в ходе интересующего исследователя возрастного промежутка называется **лонгитюдным планом**.

Следует отметить, что оппозиция меж- и внутрисубъектного плана, а также оппозиция плана поперечных срезов и лонгитюдного имеют много общего. В обоих случаях главная цель состоит в анализе результатов одних или разных людей. Относительные преимущества лонгитюдного метода и метода поперечных срезов также будут вскоре рассмотрены.

Хотя исследование Дюфресна и Кобасигавы отличалось от исследования Черри и Парка по параметру внутри/межсубъектного сравнения, оба исследования были сходны в другом, вероятно даже более принципиальном отношении. Сходство состоит в том, что оба исследования включали экспериментально изменяемую независимую переменную: легкость/трудность у Дюфресна и Кобасигавы и модель/схему у Черри и Парка. Как отмечалось в главе 2, не во всех исследованиях присутствуют такого рода настоящие независимые переменные. В так называемых корреляционных или неэкспериментальных планах исследователь просто измеряет, но не контролирует переменные и пытается выявить отношения между ними. Корреляционные планы — это третья большая тема, которая рассматривается в данной главе.

Поскольку возрастные сравнения являются центральной задачей в исследованиях в области психологии развития, глава начинается с описания планов для изучения возраста. Затем внимание переключается на методы сравнения экспериментальных условий; завершает главу анализ преимуществ и недостатков корреляционного исследования.

Возрастное сравнение

Как указывалось ранее, возраст — всего лишь одна из многих субъектных переменных, которые можно изучать. Поскольку важнейшим фактором, рассматриваемым в этой книге, является возраст, стоит отметить одно существенное свойство, отличающее возраст от большинства других субъектных переменных, — отличие, влияющее на выбор исследовательского плана. При изучении таких переменных, как пол или раса, исследователю нет необходимости ждать, когда его испытуемые перейдут с одного уровня переменной на другой; во всех исследованиях подобного рода обязательно присутствуют разные группы людей. Однако в случае с возрастом, сегодняшний шестилетка — это завтрашний 8-, или 10-, или 20-летний человек. Именно по причине этого естественного изменения по возрастному параметру

ру у изучающего возрастные различия есть выбор между внутри- и межсубъектным подходами.

Есть и еще один момент. Если мы сравниваем мальчиков и девочек, тогда, очевидно, нас интересуют различия (или, конечно, их отсутствие) между мальчиками и девочками. Если же мы сравниваем 6- и 10-летних детей, нас интересуют не только различия между 6- и 10-летними детьми, но и более глубокий вопрос. *Станут ли 6-летние дети такими как 10-летние или были ли 10-летние дети такими, как 6-летние?* Иными словами, нас волнуют не только возрастные различия, но и возрастные изменения. Как мы увидим, одной из сложнейших проблем является определение того момента, когда различия между возрастными группами действительно отражают естественные изменения, происходящие в ходе развития.

Лонгитюдный метод

В лонгитюдном исследовании одна выборка обследуется по меньшей мере дважды за определенный период времени. Хотя четких правил того, когда исследование с повторным обследованием становится «лонгитюдным», нет, при использовании этого термина учитывают по крайней мере два приблизительных критерия. Во-первых, речь обычно идет об изучении естественных, не вызванных экспериментальным путем изменений. Поэтому применение отсроченного тестирования после экспериментального вмешательства или обучения обычно не расценивается как лонгитюдное, даже если одни и те же дети проходят обследование несколько раз. Во-вторых, данный термин, как правило, употребляется в отношении регулярного обследования на довольно длительном временном интервале. Поэтому обследование одних испытуемых несколько раз в течение одной недели не должно претендовать на то, чтобы называться «лонгитюдным». Заметьте, однако, — то, что является «довольно длительным временным интервалом», зависит от уровня развития испытуемых. Недельную серию обследований можно было бы рассматривать как лонгитюдную, если на момент первого обследования испытуемым всего несколько дней от роду.

Как явствует из табл. 1.3, в обзоре журнальных статей велся учет количества лонгитюдных исследований и исследований методом поперечных срезов. Как и ожидалось, при изучении возрастных различий наиболее употребительным был метод поперечных срезов, который использовался в 57 % случаев; лонгитюдный же метод использовался в 38 % исследований, а 5 % составляли смешанные планы. Интересно отметить, что за последние годы этот диспаритет сократился. Контент-анализ тех же журналов за 1983 год дал следующие результаты: 81 % для метода поперечных срезов и только 14 % для лонгитюдного метода.

Нетрудно понять, почему лонгитюдные исследования относительно редки. На них уходит больше времени, денег и усилий. Рассмотрим в качестве примера два исследования из главы 2. На осуществление эксперимента Дюфресна и Кобасигавы потребовалось, вероятно, несколько недель. Если бы их выбор пал на лонгитюдный план, у них ушло бы минимум 6 лет. Разница даже более заметна для исследования Черри и Парка. Если бы они решились на лонгитюдный план, то им пришлось бы ждать 40–50 лет, прежде чем их молодые испытуемые превратились в пожилых.

Сами по себе широкие временные рамки лонгитюдного исследования, являются лишь практическим неудобством, конечно, досадным, но не угрожающим валидности. Со значительной продолжительностью исследования, однако, связаны дру-

гие трудности, негативно отражающиеся на валидности. Одна из них — возможность устаревания используемых методик и оборудования. Поскольку сутью лонгитюдного плана является сравнение результатов, полученных ранее, с результатами, полученными позже, исследователь вынужден продолжать использовать те способы измерения, которые выбрал в самом начале осуществления проекта. Однако зачастую, в ходе длительного исследования тест устаревает или теряет свою теоретическую актуальность; почти всегда появляются новые тесты и новые вопросы. Поэтому то, что интересно узнать в 1990 году, может отличаться от того, что было интересно узнать в 1960 году. Проблема устаревания методик особенно велика в долгосрочных исследованиях, таких как начавшиеся в 1920-х годах исследования, рассчитанные на наблюдение в течение всего срока жизни испытуемых (Kagan, 1964). Это не обязательно проблематично в относительно краткосрочных лонгитюдных планах.

Другие проблемы связаны с характером выборки. Участие в любом долгосрочном исследовании требует от испытуемых (а в случае с детской выборкой еще и от родителей испытуемых) затрат времени и усилий. Поэтому отбор, по крайней мере частично, должен основываться на таких факторах, как вера испытуемых в ценность исследования или высокая вероятность того, что они не переедут в другую местность. Но в этом случае выборка может оказаться нерепрезентативной в отношении популяции, на которую исследователь хочет перенести свои выводы. Более того, испытуемые из лонгитюдной выборки — это ровесники и составляют одно поколение или **когарту**, поэтому любые результаты в некоторой степени специфичны для данного конкретного поколения. Нас, к примеру, интересует, как изменяются люди в течение первых 30 лет жизни. Однако если все наши испытуемые родились в 1940 году, тогда с определенной долей уверенности мы сможем сказать, как изменялись люди, рожденные в 1940 году, сталкиваясь с изменениями, происходящими в мире в 1940-х, 1950-х и 1960-х. Если бы наши испытуемые появились на свет раньше или позже, мы получили бы несколько иные результаты.

Хотя лонгитюдные выборки по многим параметрам не соответствуют критерию репрезентативности, они дают возможность избежать **систематической ошибки при отборе** (Campbell & Stanley, 1966) — то есть отбора для сравнения изначально неэквивалентных групп. Если каждый испытуемый сравнивается сам с собой, никакой систематической ошибки отбора быть не может. Однако есть вероятность **избирательного отсева** (или выбывания), которое и имеет место на практике. Люди могут выходить из лонгитюдной выборки по ряду причин — в связи с переездом, нежеланием продолжать участие или (особенно в выборке пожилых людей) смертью. Если бы такие выпадения были случайными, тогда единственной неприятностью было бы сокращение размеров выборки и напрасные усилия по сбору данных, которые не будут дополнены. Однако зачастую выпадение не случайно, избирательно, то есть выбывшие из исследования испытуемые закономерно отличаются от тех, кто остался. Например, при лонгитюдном изучении интеллекта у выбывших обычно более низкий IQ при первых тестированиях (например, Siegler & Botwinick, 1979). Поскольку выпадения в связи с низким уровнем компетентности влияют на общие результаты более молодых испытуемых, следствием является «**позитивное искажение**» в пользу старших возрастных групп. Конечно, можно ограничить срав-

нение результатами людей, оставшихся в группе, и таким образом влияющими на итоговые показатели во всех возрастных категориях. Однако в этом случае изначально нерепрезентативная выборка становится еще более нерепрезентативной.

Есть и еще один признак, по которому участники лонгитюдного исследования отличаются от популяции, на которую исследователь хочет перенести свои выводы. Он очевиден: в ходе лонгитюдного исследования его участники в отличие от популяции регулярно проходят психологическое тестирование. Поэтому потенциальное значение имеют две причины искажения валидности (Campbell & Stanley, 1966). Первая — **тестирование**: влияние на результаты тестирования выполнения такого же или подобного теста прежде. Например, кажется правдоподобным, что регулярное и довольно частое выполнение теста *IQ* в конечном счете начнет отражаться на ответах испытуемого, и результат исследований действительно подтверждает это предположение (например, Nesselroade & Baltes, 1974). Вторая проблема — более общего характера — проблема реактивности. Осознание человеком того, что его изучают, может повлиять и на его поведение, что особенно выражено у участников долгосрочных лонгитюдных исследований с частым проведением измерений. Поэтому ответы этих испытуемых не всегда отражают типичный процесс развития.

Обратите внимание, что описанные выше проблемы не обязательно характерны для всех лонгитюдных исследований. Они наиболее выражены в тех из них, которые связаны с частым и явным тестированием испытуемых зрелого возраста. С другой стороны, исследователю, изучающему развитие в младенчестве методом наблюдения, нет необходимости слишком волноваться о факторе тестирования или реактивности.

Последний пункт уточняет сделанное ранее замечание о том, что лонгитюдную выборку составляют люди одного поколения. В лонгитюдном исследовании неизбежно присутствует смещение возраста и исторического времени тестирования. Это смещение является следствием проведения внутрисубъектного сравнения; если нам нужны разные возрастные группы, мы должны тестировать в разное время. Предположим, мы хотим изучить изменения, происходящие в период с 15 до 20 лет. Мы отбираем 15-летних испытуемых 1975 года рождения и обследуем их вновь в 20-летнем возрасте. Если результаты повторного тестирования отличаются от результатов первого, мы можем объяснить это двояко: тем фактом, что испытуемые повзрослели на 5 лет, или тем, что первое тестирование происходило в 1990, а второе — в 1995 году. В лонгитюдном исследовании возраст практически неотделим от времени обследования.

Насколько вероятно, что возможная проблема действительно станет проблемой? Одной из детерминант, несомненно, является характер изучаемого феномена. За примером обратимся к старшей возрастной группе. Допустим, вас интересуют изменения остроты зрения в старости. Вы тестируете выборку 60-летних лиц в 1970 году, и затем их же в возрасте 70 лет в 1980 году. Хотя историческое время логически одно из возможных объяснений любых обнаруженных изменений, в случае, когда в качестве зависимой переменной выступает острота зрения, оно не слишком правдоподобно. Что более вероятно, если вы все-таки нашли различия, это то, что в период между 60 и 70 годами зрительная система претерпевает ряд

естественных изменений. Однако представьте, что вы исследовали бы не остроту зрения, а отношение к политическим лидерам. Вы обнаруживаете, что 70-летние люди относятся к ним более негативно, чем 60-летние. Очевидный случай политического отчуждения с возрастом? Вряд ли, учитывая политические события начала 70-х. На этот раз более правдоподобным кажется историко-культуральное объяснение. Однако в любом варианте стандартный лонгитюдный план допускает выводы, в лучшем случае правдоподобные, но никак не бесспорные. Смещения возраста и исторического периода никогда не избежать.

Если учесть всю массу неприятностей, обрушивающуюся на исследователя, выбравшего лонгитюдный план, встает вопрос о том, зачем кому-то, кроме ярко выраженного мазохиста, вообще пытаться провести лонгитюдное исследование. Ответ, как и можно было ожидать, состоит в том, что лонгитюдный подход имеет ряд достоинств (Bullock, 1995; McCall, 1977). Именно к позитивной стороне мы сейчас и обратимся.

Я уже провел границу между возрастным изменением и возрастным различием. Если изучаются разные выборки разных возрастов, единственное, что можно измерить непосредственно, это возрастные различия, и лишь подразумевается, что обнаруженные различия отражают возрастные изменения. В лонгитюдных же исследованиях возрастные изменения непосредственно измеряются, а не просто подразумеваются. Как мы увидели, можно по-разному судить о причинах возрастных изменений и о том, насколько они присущи всей популяции. Но по крайней мере в центре всегда остается основной вопрос психологии развития — вопрос внутриличностного развития во времени.

Сосредоточение на внутриличностном развитии делает лонгитюдный подход единственно правильным для оценки индивидуальной стабильности или нестабильности. Предположим, вы хотите узнать, остается неизменным *IQ* ребенка в ходе его развития, повышается или снижается. Совершенно ясно, вы не можете ответить на этот вопрос, тестируя разных детей разного возраста; вы должны наблюдать одного ребенка в процессе его развития. Вопрос «стабильности *IQ*» фактически является предметом множества лонгитюдных исследований, давних (например, Thorndike, 1933) и недавних (например, Siegel, 1992). Всегда, когда в центре внимания находится индивидуальная стабильность или изменчивость, лонгитюдный подход не просто тонкость; это совершенная необходимость.

Достоинства лонгитюдного исследования не ограничиваются возможностью отслеживать развитие одной черты или одной формы поведения. Его возможности гораздо шире, поскольку оно позволяет изучить любые закономерности возрастных изменений, если только удастся измерить необходимые параметры. В одних случаях в центре внимания связь между одним аспектом развития ребенка в раннем возрасте и каким-то другим аспектом в дальнейшей жизни. К примеру, мы можем попытаться определить, связана ли скорость роста костей скелета в первые 2 года жизни с возрастом начала пубертата. В других случаях исследование направлено на выявление связи между некоторым аспектом среды, окружающей ребенка в раннем возрасте, и некоторым аспектом развития в дальнейшей жизни. Мы можем, к примеру, попытаться определить, связаны ли особенности воспитания в первые 2 года жизни с некими особенностями личности в среднем детстве или в

подростковом периоде. Всегда, когда предметом исследования выступает связь между чем-то, что имело место раньше, и чем-то, что имело место позже, лонгитюдный подход совершенно необходим.

Кроме того, лонгитюдное исследование особенно эффективно для отслеживания непрерывных, постепенных трансформаций, происходящих с поведенческими системами общего характера в процессе развития ребенка. Для прояснения этой довольно туманной фразы требуются примеры, и два примера сразу придут в голову любому, кто знаком с исследованиями в области психологии развития. Один из них — проводившиеся Пиаже исследования развития интеллекта в младенчестве (Piaget, 1952). Пиаже лонгитюдно изучал троих своих детей, с рождения и до двух лет, аккуратно фиксируя этапы развития в разных сферах интеллекта и связи между ними. Результатом явилась концепция интеллекта в младенчестве, которая по широте и глубине понимания превзошла все разработанные ранее концепции и послужила моделью для множества исследований, проводившихся в дальнейшем. Возможно, что по крайней мере некоторые из тех же открытий могли бы быть сделаны и с продуманным использованием метода поперечных срезов на выборке малышей разных возрастов; однако сомнительно, чтобы без тщательного, практически ежедневного изучения изменений, происходящих с одним ребенком, удалось бы составить полное представление об интеллекте в период младенчества.

То же можно сказать и об исследованиях раннего речевого развития (например, Brown, 1973). Так же как и Пиаже, исследователи использовали лонгитюдный подход для изучения постепенных изменений, происходящих в первые годы освоения речи. Что, к примеру, представляет из себя самая ранняя форма отрицания, и как эта форма в конечном счете превращается в сложную систему правил, которые применяют старшие дети и взрослые? Вновь тщательное лонгитюдное изучение, при котором регистрируются изменения, происходящие у одного ребенка, позволяет получить представление о речи в раннем возрасте и о ее развитии, представление, которое, по всей видимости, не удалось бы получить, используя только метод поперечных срезов.

Ясно, что лонгитюдное исследование такого рода подразумевает нечто большее, чем просто повторное тестирование одного и того же ребенка; это скорее жизнеописание. В каких случаях тщательное лонгитюдное изучение приносит наибольшие плоды? Замечу, во-первых, что оно наиболее подходит для младших детей, которые не склонны изменять свое поведение, находясь под наблюдением. Если отбросить этот практический момент, несомненно, главным доводом в пользу лонгитюдного метода описанного выше типа является возможность его применения при проведении новаторских исследований, в которых еще только предстоит открыть множество существенных феноменов. Слово «новаторство» как нельзя лучше подходит для описания исследований Пиаже, направленных на изучение младенческого интеллекта. Как только появляется некое представление об общем направлении развития и о поворотных пунктах этого развития, с пользой может применяться более прицельное изучение методом поперечных срезов, кроме того, лонгитюдное исследование особенно подходит для отслеживания процесса постепенного формирования новых способностей, медленного прохождения разнообразных промежуточных этапов на пути к зрелости. Как (если дополнить пример с от-

рицанием, проблематикой, которую изучал Пиаже) элементарный хватательный рефлекс становится целенаправленным, зрительно координируемым актом хватания? Наконец, тщательное, долгосрочное изучение одних и тех же детей может оказаться незаменимым в интерпретации поведения, то есть в попытках перейти от внешних границ поведения к представлению о том, что лежит в его основе (когнитивная структура, правило языка, личная реакция или что-то еще). В большинстве случаев исследователь видит испытуемых первый и единственный раз, когда они приходят для тестирования, и его способность понять смысл их поведения зависит от этого непродолжительного периода взаимодействия. Пиаже, напротив, изучал детей буквально с рождения, а его полная осведомленность о жизни и особенностях каждого ребенка создавала прочную основу для интерпретации любой формы поведения¹.

Последний аргумент в пользу лонгитюдного подхода носит негативный характер. Главной альтернативой лонгитюдного метода является метод поперечных срезов, который также имеет ряд недостатков. Проблемы, которые могут возникнуть в исследованиях методом поперечных срезов, являются темой следующего раздела.

Метод поперечных срезов

Метод поперечных срезов предполагает тестирование разных людей разного возраста. По этой причине этим методом нельзя измерить непосредственно возрастные изменения, невозможно с его помощью и ответить на вопрос об индивидуальной стабильности во времени. Как мы видели, эти ограничения являются одним из оснований для использования лонгитюдного плана.

Но есть и другие недостатки. В связи с изучением разных выборок и разных возрастных групп возникает вероятность *систематической ошибки при отборе*. Возможно, сравниваемые группы различаются не только по изучаемой независимой переменной (в данном случае по возрасту), но и по другим параметрам, и именно эти последние различия обуславливают различия в значении зависимой переменной. Мы уже касались этого вопроса в главе 2, рассматривая особенности возраста как независимой переменной. Как отмечалось, цель состоит не в исключении всех межгрупповых различий, за исключением возрастных, а только тех, которые не имеют естественной связи с возрастом. Отмечалось также, что в большинстве случаев ответ на вопрос, какие параметры следует уравнивать, довольно очевиден — например, пол, расу, социальный класс, IQ. Однако теперь следует добавить, что на практике добиться желаемого соответствия не всегда просто. Обычно исследователи в области психологии развития отбирают испытуемых разных возрастов, используя разнообразные источники: новорожденных — из роддомов; младенцев — из семей, откликнувшихся на просьбу принять участие в исследовании; дошкольников — из детских садов; детей 5–11 лет — из начальной школы; подростков — из школ и колледжей. Эти различия в обстановке могут создавать различия и между популяциями. Поэтому, хотя исследователь может осознавать значение приведения в соответствие ряда параметров, сформировать действительно равноценные группы может оказаться довольно трудно.

¹ Гинсбург и Оппер (Ginsburg & Oppen, 1988) обсуждают и этот момент, и другие достоинства подхода Пиаже.

Систематическая ошибка может также принимать форму *избирательного отсева*. Изначальная эквивалентность групп может растаять на глазах, если часть испытуемых откажется от продолжения обследования до его завершения. Проблема не только в том, что в одной возрастной группе оказывается больше выпадений, чем в другой. Суть проблемы та же, что и в лонгитюдном исследовании: те, кто выбывает, нередко отличаются от тех, кто остается. Поэтому валидности угрожает именно «избирательность» избирательных выпаждений.

Нетрудно представить себе ситуацию, когда избирательное выпадение может исказить результаты сравнения возрастных групп. Предположим, что мы изучаем дошкольников и делим нашу выборку на младших (2,5–4 года) и старших (4–5,5 лет) детей, получая при этом две группы для сравнения. Процедура тестирования довольно утомительна, она требует от ребенка осмысления ряда инструкций и ответов на вопросы в течение весьма длительного периода времени. Не все дошкольники к этому готовы, и некоторые поэтому выбывают из исследования. Наибольшие шансы на выбывание у детей младшей группы. Наибольшие шансы на выбывание также у тех, кто обладает наименьшей компетентностью среди всех испытуемых из выборки. В таком случае мы останемся с двумя неравноценными группами: достаточно репрезентативной выборкой старших детей и явно нерепрезентативной по параметру уровня способностей выборкой младших детей. Очевидно, что любое такое выпадение снизило бы вероятность выявления улучшения результатов с возрастом.

Вернемся к вопросу о первичном отборе испытуемых. Я уже дважды говорил о том, что, как правило, выбор параметров для приведения в соответствие не представляет труда. Пора теперь обратиться к исключениям из этого «правила». Сомнения в том, по каким показателям следует уравнивать, чаще всего возникают в случаях, когда между группами очень большая разница в возрасте, и поэтому потенциально существует множество других различий. Поэтому эти сомнения наиболее сильны в ситуациях сравнения выборки пожилых с выборкой молодых людей. Наглядный пример — переменная образовательного уровня: сегодня процент лиц, окончивших среднюю школу, значительно выше, чем несколько десятилетий назад. Допустим, нам нужно сравнить 25-летних и 75-летних испытуемых. Если в обе возрастные группы мы отбираем по принципу случайности, молодые испытуемые в среднем окажутся образованнее пожилых. Тогда мы получим смешение возраста и уровня образования. Если мы включим в выборку пожилых людей только тех, у кого выше образовательный уровень, то получим равноценные выборки по параметру уровня образования, но за счет нерепрезентативности и искажения выборки пожилых людей в положительную сторону. Ни то, ни другое нежелательно; вероятно, если возможно, лучшим вариантом в этом случае будет совмещение двух подходов (см. Green, 1969). Однако главное это то, что при любой попытке сравнить взрослых испытуемых разных возрастов на определенном историческом этапе неизбежно происходит смешение возраста и уровня образования.

Сказанное выше о приведении в соответствие параметров группы фактически является частным случаем особенности, характерной для всех исследований методом поперечных срезов. Ранее отмечалось, что лонгитюдный подход к изучению возрастных различий неизменно связан со смешением возраста и времени обследования. Добавим теперь, что использование метода поперечных срезов неизбеж-

но связано со смешением возраста и поколения, или когорты. Поскольку испытуемые входят в разные возрастные группы, они появились на свет в разное время и росли в разных условиях. Один из примеров такого межпоколенного различия — несопоставимость образовательных возможностей, которые были у сегодняшних 25-летних к сегодняшним 75-летним людям. Можно привести еще массу примеров. Сегодняшние 75-летние люди, будучи детьми и подростками, пережили Великую Депрессию, их молодость совпала со Второй мировой войной, они достаточно долго жили без телевидения и других атрибутов современной жизни и т. д. Предположим теперь, что мы обнаружили различия в значении зависимой переменной между группой 25-летних и группой 75-летних лиц. Следует нам относить это различие на счет возраста или на счет разницы между поколениями?

Так же, как и в случае причин искажения валидности, рассмотренных в этой главе, степень смешения возраста и когорты зависит от характера проводимого исследования. Оценивая вероятность эффекта поколения, важно учитывать два фактора. Первый — тип изучаемой зависимой переменной. Если нас интересуют политические взгляды или результаты выполнения тестов интеллекта, тогда эффект поколения может быть достаточно выражен; практически, этот эффект был наглядно продемонстрирован в исследовании *IQ* (например, Schaie, 1994). Если нас интересуют изменения в частоте сердечных сокращений или остроте зрения, тогда эффект поколения, скорее всего, будет играть незначительную роль. В целом, чем более «универсальный», более «биологический» характер носит зависимая переменная, тем меньше вероятность того, что ее величина будет разной в разных когортах. Однако помните, что почти всегда возникают споры о том, насколько «универсальна», неспецифична конкретная переменная. Возможно, к примеру, что острота зрения изменяется от поколения к поколению как функция от изменения таких факторов, как уровень искусственной освещенности, наличие или отсутствие телевидения в годы становления индивида.

Другой фактор, который следует учитывать, — разница в возрасте. Эффект поколения становится особенно серьезной проблемой в исследовании групп с большой разницей в возрасте. В действительности, впервые вопрос межпоколенных различий встал в исследовании, направленном на сравнение выборок молодых и пожилых людей, и до сих пор наиболее часто он обсуждается именно в этом контексте. С другой стороны, психологу, сравнивающему 3- и 4-летних детей, нет необходимости беспокоиться о том факте, что одни дети родились в 1990 году, а другие — в 1991. В границах детского возраста испытуемых обычно рассматривают как представителей одной когорты. Однако и здесь могут возникнуть сомнения. Что если сравнить когорту, выросшую без «Улицы Сезам», с когортой, выросшей с «Улицей Сезам»? Что если сравнить 11-летних детей, которые в начальной школе обучались по новому методу, с 7-летними, которых стали обучать по старой схеме? Мы живем во времена стремительных изменений в культуре и образовании, и эти изменения могут отражаться на результатах сравнения по крайней мере некоторых возрастных групп детей.

Последняя проблема, о которой стоит упомянуть, — это проблема **эквивалентности измерений**. Для сравнения уровня выраженности определенного поведения или определенной способности у разных возрастных групп необходима процедура

точной оценки этой формы поведения или способности для каждой изучаемой группы. Однако зачастую тест, подходящий для одного возраста, не подходит для другого. Например, тест способности к классификации может быть точным индикатором уровня развития этих навыков у 7-летних детей, однако, с точки зрения речевых навыков, слишком сложным для многих 4-леток. Тогда в разных возрастных группах он будет измерять разные навыки: навыки классификации у 7-летних детей и объем словарного запаса у 4-летних. Заметьте, что тест все еще вскрывал бы истинные и довольно значимые возрастные различия: 7-летние дети по этому показателю действительно превосходят 4-летних, однако основанием для различий служит не то, что стремился оценить исследователь.

Проблема эквивалентности измерения выходит за рамки метода поперечных срезов. Она встает всякий раз при сравнении разных возрастных групп, поэтому равно актуальна и для лонгитюдного исследования. Однако форма, которую она принимает в лонгитюдном исследовании, несколько иная. Возьмем, к примеру, лонгитюдное изучение агрессии (например, Cairns, Neckerman, Ferguson & Gariepy, 1989). Исследователя, изучающего агрессию у детей сначала 4-летнего, а затем 12-летнего возраста, вряд ли будет интересовать сравнение выраженности агрессии у этих возрастных групп. Если бы в центре внимания находился уровень агрессии, тогда возникли бы серьезные осложнения, обусловленные тем, что формы, которые принимает агрессия, и обстоятельства, провоцирующие ее в 12-летнем возрасте, не те же, что в 4-летнем. Лонгитюдное изучение детей означает, что в центре внимания скорее всего будет находиться *стабильность индивидуальных различий* по параметру агрессии в процессе развития детей. Иными словами, вопрос состоит в следующем: остаются ли более (менее) агрессивными в 12-летнем возрасте дети, которые были более (менее) агрессивными в 4-летнем возрасте? Ребенок может быть крайне агрессивен и в 4 года, и в 12 лет, хотя частота и формы проявления агрессии изменяются. Этот акцент на относительном положении в группе, а не на абсолютном показателе, частично разрешает проблему эквивалентности измерения. Однако следует помнить, что в обеих возрастных группах необходимо использовать валидные методики измерения.

Более сложные планы

Из сказанного выше следует, что и лонгитюдный метод, и метод поперечных срезов имеют ряд недостатков. Таблица 3.1 наглядно представляет рассмотренные ранее проблемы. Некоторые из них, по крайней мере в принципе, разрешимы — например, систематическая ошибка при отборе в исследовании методом поперечных срезов. Однако другие недостатки присущи как лонгитюдному плану, так и плану поперечных срезов, и по этой причине их никак нельзя преодолеть. В частности, в исследовании методом поперечных срезов нельзя избежать смешения возраста и поколения, а в лонгитюдном исследовании — возраста и времени измерения.

В последние годы идет активное обсуждение недостатков традиционных лонгитюдных и поперечных планов, которое стимулировало разработку рода новых процедур для измерения возрастных изменений. Поскольку эти процедуры пока применяются в основном в исследованиях лиц пожилого возраста, их рассмотре-

ние отложено до главы, посвященной старению (глава 13). Однако здесь можно сделать несколько предварительных замечаний.

Рисунок 3.1 схематически иллюстрирует планы, описанные выше. В таблице представлены возрасты, соответствующие определенным годам рождения и измерения. Лонгитюдный план соответствует любой строке таблицы. В этом случае выборка испытуемых, рожденных в одном году, регулярно обследуется в течение некоторого периода времени. Поперечный план соответствует любому столбцу. В этом случае отдельные выборки испытуемых разных возрастов обследуются в одно время.

Таблица 3.1

Недостатки лонгитюдного и поперечного планов

Лонгитюдный план	План поперечных срезов
Практические трудности (временные и денежные затраты)	Косвенное измерение возрастных изменений
Вероятность устаревания методик	Невозможность применения для изучения индивидуальной стабильности
Вероятность нерепрезентативности выборки	Вероятность систематической ошибки при отборе
Ограничение рамками одной когорты	Вероятность избирательного выпадения
Вероятность избирательного выпадения	Трудность в установлении эквивалентных критериев
Эффект повторного тестирования	Смещение возраста и времени рождения (когорты)
Трудность в установлении эквивалентных критериев	
Смещение возраста и времени измерения	

На рис. 3.1 изображен также еще один, не обсуждавшийся ранее вид плана — **план с временным лагом**. Его представляют диагонали таблицы. Мы могли бы изучать выборки 40-летних лиц в 1960 году, другую выборку 40-летних лиц в 1990 году, еще одну выборку 40-летних лиц в 2000 и еще одну — в 2010 году. Очевидно, что план с временным лагом не дает информации непосредственно о возрастных изменениях или возрастных различиях, поскольку обследуется только одна возрастная группа. Однако он может дать информацию о факторах, приводящих к смещению в лонгитюдных исследованиях и исследованиях методом поперечных срезов. В частности, обнаружив различия между нашими выборками 40-летних испытуемых, мы поймем, что эти различия отражают либо фактор когорты (главный источник смещения в поперечном плане), либо фактор времени тестирования (главный источник смещения в лонгитюдном плане), либо, естественно, комбинацию этих двух факторов. Невозможность определения, какой из факторов имеет первостепенное значение, свидетельствует о том, что план с временным лагом приносит свой собственный вид смещения — между когортой и временем тестирования.

Этот тип исследования используется довольно редко. Иногда сравнение с временным лагом возможно просто по причине естественного исторического хода науки. Например, Пиаже начал изучать понимание детьми принципа сохранения в 1930–1940-е годы. Следствием прихода популярности к его концепции несколько десятков лет спустя явилась вторая волна изучения понимания принципа со-

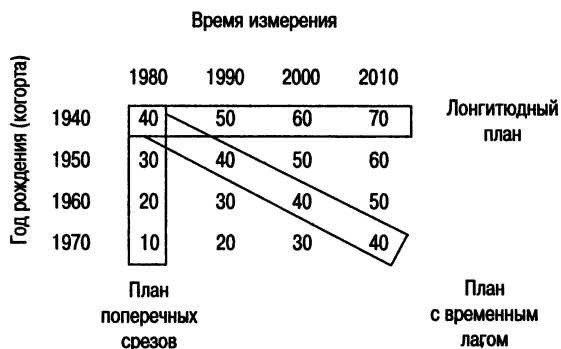


Рис. 3.1. Примеры лонгитюдного плана, плана поперечных срезов и плана с временным лагом (числа, составляющие массив таблицы, соответствуют возрасту)

хранения, которая пришлась на 1960-е 1970-е годы. Вместе ранние и поздние исследования составляют план с временным лагом: две группы детей одного возраста, но появившихся на свет и проходивших тестирование в разное время. В этом случае не имело значения ни поколение, ни время тестирования, поскольку дети 1970-х выполняли задания на сохранение точно так же, как и дети 1930-х. Оценка уровня интеллекта дала совершенно иную картину. Результаты выполнения ребенком *IQ*-теста сравниваются с результатами детей того же возраста из нормативной выборки, которая набирается при создании теста, о чем более подробно будет рассказано в главе 11. Поэтому, скажем, 10-летний ребенок, выполняющий тест *IQ* в 1995 году, сравнивается с 10-летним ребенком, выполнявшим этот тест в 1975 году. Как правило, в этих сравнениях обнаруживается тенденция к улучшению результатов со временем (что, в действительности, является основанием для периодической перенормировки тестов интеллекта). Поскольку возраст остается неизменным, мы знаем, что улучшение должно отражать эффект либо времени измерения, либо (что более вероятно) когорты.

Лонгитюдный план, а также план поперечных срезов и план с временным лагом иногда называют «простыми». Они просты по сравнению с альтернативным, определенно *не* простым последовательным планом. **Последовательный план** соединяет в себе элементы лонгитюдного и поперечного планов и плана с временным лагом. Целью последовательного плана является разграничение эффектов возраста, поколения и времени измерения. Эти элементы могут по-разному сочетаться и анализироваться, отсюда несколько типов последовательных планов. В этой главе кратко и в общих чертах описано два типа. В главе 13 мы вернемся к теме последовательных планов и рассмотрим их конкретно, а не гипотетически, а также обсудим актуальную на сегодняшний день схему использования таких планов.

Сначала несколько слов о логике последовательных планов. В идеальном варианте нам бы хотелось иметь возможность оценить вклад каждого из трех потенциально значимых факторов — возраста, поколения и времени тестирования — в рамках одного аналитического метода. К сожалению, этому мешает их взаимозависимость; как только установлены любые два из них, автоматически устанавливаются и уровни третьего. Решив, к примеру, изучать определенные возрастные группы и

Год рождения (когорты)	Время измерения			
	1980	1990	2000	2010
1940	40	50	60	
1950		40	50	60

Рис. 3.2. Пример плана последовательности когорт (числа в таблице соответствуют возрасту)

Год рождения (когорты)	Время измерения		
	1990	2000	2010
1930	60		
1940	50	60	
1950	40	50	60
1960		40	50
1970			40

Рис. 3.3. Пример плана последовательности времени (числа в таблице соответствуют возрасту)

определенные когорты, мы тем самым с необходимостью фиксируем время измерения, определяемое нужным сочетанием возраста и когорты. Следствием этой взаимозависимости является то, что в рамках одного аналитического метода в качестве независимых переменных могут выступать только два из трех факторов. В разных последовательных планах акцентируется внимание на разных факторах. В первом из приведенных ниже примеров независимыми переменными являются возраст и когорта, во втором — возраст и время измерения.

Рисунок 3.2 иллюстрирует *план последовательности когорт*, при котором отбираются и регулярно обследуются в одном и том же возрасте равное количество лет испытуемые разных когорт. Таким образом, этот план состоит из двух (или более) накладывающихся друг на друга лонгитюдных исследований. В примере, который иллюстрирует рисунок, группы испытуемых, родившихся в 1940 и в 1950 годах, обследуются по три раза на протяжении 20-летнего периода. Такой план обладает рядом преимуществ перед стандартным лонгитюдным исследованием или исследованием методом поперечных срезов, поскольку: а) измерения происходят в разное время, переменная возраста не смешивается с переменной когорты (главное смещение в поперечном плане); б) выборки состоят из испытуемых, родившихся в разное время, лонгитюдное сравнение не ограничивается рамками одной когорты или одного поколения; в) разные возрастные группы обследуются одновременно, измерение происходит как в лонгитюдном, так и в поперечном направлении; г) одна возрастная группа обследуется в разное время, измерение происходит в направлении временного лага. Иными словами, этот план дает больше информации, чем любой из стандартных, и, таким образом, больше шансов вычленить вклад каждого из факторов.

Рисунок 3.3 иллюстрирует *план последовательности времени*, который включает в себя два (или более) исследования методом поперечных срезов, производимых в разное время. В этом примере, в 1990, 2000 и 2010 годах сравниваются выборки 40-, 50- и 60-летних испытуемых. При обследовании в разное время испытуемые могут быть либо независимыми (то есть новые люди при каждом обследовании), либо те же (если первичная выборка изучается в лонгитюде). Этот план имеет то же достоинство, что и план последовательности когорт — он дает больше информации, чем простые планы. Его особым преимуществом является возможность

разграничения переменных возраста и времени тестирования (главное смещение в лонгитюдном исследовании).

Если в разное время изучаются независимые выборки, удастся избежать и некоторых проблем, связанных с лонгитюдным методом (избирательных выпадений, эффектов повторного тестирования).

Как уже говорилось, более подробно об этих планах можно узнать из главы 13. Однако здесь следует сделать несколько замечаний. Во-первых, очевидно, что последовательные планы, хотя и более информативны, требуют больше затрат — времени, усилий и денег, — чем простые поперечные и лонгитюдные планы. На реализацию плана, представленного на рис. 3.3, к примеру, ушло бы 20 лет, при этом нужно было бы сформировать (в случае использования независимых выборок) 9 групп испытуемых. Любой исследовательский проект предполагает выполнение ряда задач, только часть которых может быть выполнена на практике. Лучшими являются те планы, которые действительно можно реализовать.

Второе замечание касается критики, которой подвергаются традиционные лонгитюдный и поперечный планы. В некоторых случаях эта критика довольно сурова. Болтс, Рис и Нессельроде (Baltès, Reese, & Nesselrode, 1977), к примеру, пишут, что «такие простые методы, как лонгитюдный и метод поперечных срезов, демонстрируют настолько низкий уровень контроля, что данные, полученные при помощи любого из них, по большей части невалидны и бесполезны для исследователя» (р. 124). Это серьезное обвинение, учитывая, что (по скромным оценкам) 99 % имеющихся у нас данных о возрастных изменениях получены в исследованиях методом поперечных срезов и в лонгитюдных исследованиях. Действительно ли почти все из того, что было сделано в психологии развития, бессмысленно?

Многие исследователи оспорили бы столь пессимистический вердикт. Действительно, цель качественного исследования заключается в том, чтобы минимизировать вероятность искажения валидности, и в идеале план должен исключать эту вероятность. К сожалению, никогда нельзя быть уверенными в отсутствии альтернативной интерпретации результатов. Суть в том, насколько правдоподобны эти альтернативные интерпретации. Особенно когда речь идет о выборке детей и особенно когда изучаются базовые формы развития (физическое созревание, константность восприятия; феномены, описанные Пиаже, языковые правила — список можно продолжить), использование традиционных лонгитюдного метода и метода поперечных срезов вполне оправданно. Предположим, что мы изучаем концепцию Пиаже о принципе сохранения числа — то есть понимания того, что число объектов неизменно, несмотря на изменение их наложения, приводящее к кажущемуся противоречию, мы обнаруживаем, что 6-летние дети осознают сохранение числа, а 4-летние — нет. Крайне маловероятно, что этот результат является следствием факта рождения одних детей в 1988 году, а других — в 1990 году (если это исследование методом поперечных срезов) или фактом тестирования одной группы в 1992 году, а другой — в 1994 году (если это лонгитюдное исследование). Гораздо более правдоподобно (а на современном уровне понимания фактически бесспорно) то, что эти результаты отражают истинные возрастные изменения. В таких случаях при использовании метода поперечных срезов или лонгитюдного метода мы полу-

чаем данные достаточно валидные и достаточно ценные. (С другими аргументами в защиту традиционных методов, в особенности лонгитюдного, можно ознакомиться у McCall, 1977.)

Сравнение условий

Внутрисубъектные/межсубъектные планы

Обратимся теперь к вопросу о том, как сравнивать два и более видов тестовых или экспериментальных условий. Как уже отмечалось, в принципе, возможно два подхода: либо дать одним и тем же испытуемым все задания (поставить их во все экспериментальные условия), либо приписать разных испытуемых к разным экспериментальным группам. Первое называется *внутрисубъектным планом*, последнее — *межсубъектным планом*. Поскольку рассмотрение этих двух подходов потребует постоянных перекрестных ссылок и сравнений, проще всего описывать их параллельно, а не по отдельности.

Как исследователю решить, осуществлять сравнение внутри группы или между группами? Как и в случае с лонгитюдным методом и методом поперечных срезов, зачастую определенную роль играет вопрос удобства. Обычно (когда конкретно, мы вскоре уточним) для внутрисубъектного исследования требуется меньше испытуемых. Предположим, нам нужно сравнить уровень трудности трех заданий, и мы знаем, что для выявления различий в трудности нужно по меньшей мере 20 испытуемых. Всегда, когда контингент потенциальных испытуемых ограничен, экономичность внутрисубъектного плана представляется довольно заманчивой.

Однако соображения удобства не всегда склоняют чашу весов в пользу внутрисубъектного подхода. За меньший размер выборки во внутрисубъектном исследовании приходится платить — конкретно необходимостью больше времени уделять каждому испытуемому либо за счет более длительных обследований, либо за счет большего их числа. Длительные или частые обследования подвергают серьезному испытанию терпение и мотивацию ребенка, особенно маленького, даже если исследователя не тревожит возможность того, что от ребенка требуется слишком много, это может беспокоить родителей или школьную администрацию. В таких случаях межсубъектный план, в котором требования к одному ребенку сведены к минимуму, является наиболее разумным выбором.

Статистические соображения также могут влиять на решение в пользу либо внутри-, либо межсубъектного плана. Статистические критерии, используемые для сравнения внутри группы, несколько отличаются от тех, которые используются для межгруппового сравнения. Более того, внутрисубъектные критерии зачастую обладают большей мощностью, чем межсубъектные — то есть скорее выявят значимое различие, если оно действительно существует. Большая мощность этих критериев обусловлена уменьшением нежелательной дисперсии, которое обеспечивает внутрисубъектный план. Вспомним обсуждавшиеся ранее понятия первичной дисперсии и вторичной дисперсии, или дисперсии ошибки. Как отмечалось, целью качественного экспериментального плана является максимизация первичной дисперсии и минимизации нежелательной дисперсии, источником которой являются

другие факторы. Было также отмечено, что одним из источников нежелательной дисперсии являются индивидуальные различия между испытуемыми. Обследование одних и тех же испытуемых в разных экспериментальных условиях позволяет снизить эту дисперсию и таким образом расширить возможности сравнения. Как следствие, повышается вероятность того, что показатель различия достигнет уровня статистической значимости.

И внутри-, и межсубъектный планы подвержены особым формам систематических ошибок. Очевидным недостатком межсубъектного плана является возможность и систематической ошибки отбора. Поскольку разные люди ставятся в разные условия, всегда существует вероятность того, что любые обнаруженные различия отражают не истинный эффект экспериментальных манипуляций, а изначальные индивидуальные различия между испытуемыми. Во внутрисубъектном плане, в котором каждый испытуемый обследуется при каждом из условий, эта вероятность отсутствует. Заметьте, что данное преимущество внутрисубъектного плана перед межсубъектным соотносится с ранее описанным преимуществом лонгитюдного метода перед методом поперечных срезов.

Существует два способа исключить возможность систематической ошибки отбора в межсубъектном исследовании (вспомните табл. 2.3). Первый — уравнивать потенциально значимые переменные; ниже мы рассмотрим плюсы и минусы уравнивания. Другой подход был описан в главе 2: случайное причисление испытуемых к разным группам. Если размер выборки достаточно велик, если причисление действительно производится по принципу случайности, изначальные различия между испытуемыми будут находиться под контролем, а смещения субъектных переменных и условий удастся избежать. Как указывалось в главе 2, логика подхода, основанного на случайности, безупречна; проблема в том, чтобы обеспечить положительный ответ на оба «если».

Наиболее явная причина искажения валидности в рамках внутрисубъектных планов — вероятность эффектов повторного тестирования. Рассмотрим исследование, направленное на сравнение уровня трудности нескольких когнитивных заданий. Исследователь выбрал внутрисубъектный план, в котором каждый ребенок выполняет все задания. Поскольку на выполнение каждого из них требуется время, в процессе решения дети могут утомиться и потерять интерес. Поэтому успешность выполнения последних заданий может оказаться ниже успешности выполнения первых заданий. Или наоборот, в начале исследования дети могут испытывать некоторую робость и смущение, но освоиться в процессе тестирования. В этом случае успешность будет повышаться. В обоих примерах эффект повторного тестирования исказил бы результат сравнения трудности заданий, который собственно и интересует исследователя.

Описанные выше эффекты «вработывания» и «утомления» попадают в разряд эффектов последовательности. Термин **эффект последовательности** характеризует любую общую тенденцию закономерного изменения успешности выполнения заданий от первых к последним. Обычно закономерное изменение заключается либо в общем повышении, либо в общем снижении успешности. Другая потенциальная проблема, связанная с внутрисубъектными планами, — это вероятность эффекта переноса. Об **эффекте переноса** говорят, когда ответ на одно задание (ре-

акция в одних условиях) изменяется в зависимости от того, предшествует ему некое другое задание (другие условия) или следует за ним. Приведем пример, поясняющий это кажущееся довольно туманным определение. Представьте, что нужно сравнить уровень трудности двух заданий: *A* и *B*. Допустим, что по отдельности эти задания могут правильно решить 50 % испытуемых. Однако оказывается, что если сначала идет задача *A*, ее решение подсказывает путь решения задачи *B*, в результате частота правильных ответов на задачу *B* повышается до 70 %. Выполнение же задания *B*, если оно идет первым, подсказывает испытуемому пути решения, непригодные для решения задачи *A*; в результате частота правильных ответов на задание *A* падает до 30 %. Заметьте, что в этом случае в ходе экспериментального сеанса не обнаруживается ни общего улучшения, ни общего ухудшения; суть в том, что успешность выполнения одного задания зависит от того, идет оно до или после некоего другого задания. Хотя конкретные механизмы могут различаться, смысл эффекта последовательности и эффекта переноса общий: затруднение интерпретации при сравнении заданий или условий.

Чаще всего эффект последовательности создает проблемы в тех случаях, когда экспериментатор постоянно предъявляет задания в одном порядке. Отсюда рекомендация: при сравнении заданий или условий следует избегать одинакового порядка предъявления. Есть две альтернативы неизменной последовательности. Первая — рандомизация порядка заданий или условий. В определенных случаях, особенно когда количество заданий велико, рандомизация может оказаться весьма разумным решением. Однако зачастую рандомизации следует предпочесть **контрбалансировку** последовательности предъявления. Суть контрбалансировки легче объяснить на примере, чем через определение; простой пример представлен в левой верхней части табл. 3.2. Как можно заметить, контрбалансировка — это метод такого упорядочения заданий, при котором некоторое задание занимает разные позиции в ряду равное количество раз. Так, в данном примере задание *A* появляется одинаково часто на первом, втором и третьем месте; более того, оно равно часто предшествует заданиям *B* и *C* и следует за ними. В этом случае контрбалансировка полная — то есть использованы все возможные комбинации трех заданий. Очевидно, что с увеличением количества заданий возрастает и число возможных перестановок; 4 задания можно скомбинировать 24 способами (они представлены в верхней правой части табл. 3.2), а 5 заданий — 120 способами. В таких случаях полная контрбалансировка нецелесообразна; однако все еще можно отобрать такие последовательности, которые создавали бы достаточную сбалансированность. Примеры таких последовательностей для четырех и пяти заданий представлены в нижней части табл. 3.2.

Контрбалансировка обладает двумя преимуществами перед рандомизацией. Во-первых, она исключает вероятность смешения содержания задания с порядком предъявления, чего не может сделать рандомизация. Во-вторых, поскольку смешение устраняется, исследователь получает возможность сравнить разные последовательности предъявления и выделить любые эффекты последовательности или переноса, которые могут присутствовать в данных. Однако обратите внимание на то, что эти эффекты можно обнаружить, только если выборка достаточно велика и каждый вид последовательности представлен достаточно часто. Это замечание

Таблица 3.2

Примеры полной и частичной контрбалансировки

Полная контрбалансировка	Три задания	Четыре задания			
	ABC ACB BAC BCA CAB CBA	ABCD ABDC ACBD ACDB ADBC ADCB	BACD BADC BCAD BCDA BDAC BDCA	CABD CADB CBAD CBDA CDAB CDBA	DABC DACB DBAC DBCA DCAB DCBA
Частичная контрбалансировка	Четыре задания		Пять заданий		
	ABCD BDAC CADB DCBA		ABCDE BEDCA CAEBD DCBEA EDACB		

уточняет сделанное ранее заявление о том, что для внутрисубъектного исследования обычно требуется меньше испытуемых, чем для межсубъектного: всегда, когда имеет значение возможный эффект последовательности, количество испытуемых, необходимое для внутрисубъектного исследования, существенно возрастает.

До этого момента мы рассматривали ряд факторов, которые должен учесть исследователь, делая выбор между внутри- и межсубъектным планом. Однако в некоторых случаях выбора просто нет; характер исследуемого вопроса сам диктует решение, каким должен быть план. В частности, всегда когда нужно выявить закономерности успешности выполнения заданий внутри группы, необходим внутрисубъектный план. Когда же нужно выявить определенные стойкие изменения как результат экспериментального воздействия, необходим межсубъектный план. Конкретизируем оба этих положения.

Замечание, касающееся выявления внутригрупповых закономерностей, перекликается с высказанным ранее аргументом в пользу лонгитюдных планов. Там мы указали на то, что везде, где речь идет об индивидуальной стабильности или индивидуальных изменениях во времени, требуется лонгитюдный подход, при котором испытуемые изучаются в процессе своего развития. Аналогично, везде, где речь идет о связи между двумя или более показателями в определенный момент времени, требуется внутрисубъектный подход, при котором одни и те же испытуемые оцениваются по ряду параметров. Предположим, что нам интересно, связано ли умение ребенка встать на чужую позицию с его способностью сообщать информацию другим людям (например, Shantz, 1975). Очевидно, что мы не можем оценить умение занимать чужую позицию у одной группы детей, а способность сообщать информацию — у другой, мы должны измерить оба параметра у всех детей. Предположим (возвращаясь к прежнему примеру), что мы хотим узнать, связаны ли показатели *IQ* детей с их отметками в школе. Мы не можем оценить *IQ* в одной вы-

борке, а IQ — в другой; вновь нужно измерить оба параметра у всех детей. Или допустим (предваряя пример, который более подробно разбирается в главе 11), что нас интересует последовательность овладения рядом когнитивных навыков. Действительно ли, к примеру, как утверждал Пиаже, дети осваивают принцип сохранения массы до принципа сохранения веса, а принцип сохранения веса — до принципа сохранения объема? Единственный способ узнать это — проанализировать представления детей о всех трех принципах. Эти примеры иллюстрируют основной мотив использования внутрисубъектного подхода: выявление взаимосвязей и закономерностей развития.

Замечание, касающееся экспериментального воздействия, приводящего к изменениям, в некотором отношении сходно с тем, что говорилось ранее об эффекте тестирования в лонгитюдных планах и эффекте переноса во внутрисубъектных планах. Суть в том, что опыт выполнения определенного задания или пребывания в определенных экспериментальных условиях может сделать испытуемого непригодным для оценки выполнения им других заданий или реакции в других экспериментальных условиях. Допустим, мы хотим сравнить эффективность нескольких методов обучения принципам сохранения (например, Smith, 1968). Мы набираем группу детей, не имеющих представлений о сохранении, и ставим их в условия обучения *A*. Вряд ли имеет смысл ставить затем этих же детей в условия *B*, так как, если условия *A* оказались эффективными, многие из детей уже будут владеть принципом сохранения! Тот же довод применим и в отношении любого исследования, цель которых добиться стойких изменений у испытуемых — программы вмешательства для так называемых неблагополучных детей, терапевтические программы для детей с психическими нарушениями, образовательные программы для родителей, ожидающих ребенка, и т. д. Во всех случаях, желая сравнить эффективность разных программ, мы должны использовать межсубъектный план, при котором разные испытуемые ставятся в разные экспериментальные условия. Обратите внимание также, что справедливость данного довода не ограничивается ситуациями активного изменения (вмешательство или терапия); он справедлив и для более локальных, краткосрочных изменений. Предположим, что нас интересует вопрос, помогает ли детям обучение проговаривать запоминаемое вслух при выполнении заданий на кратковременную память (например, Ferguson & Gray, 1976). Вряд ли дети, которых научили этой стратегии, перестанут ее использовать, как только мы перестанем давать соответствующую инструкцию; желая выявить эффект проговаривания, мы должны протестировать разные группы испытуемых. По поводу последнего примера и сделанного на его основе вывода может возникнуть возражение. В случае с проговариванием нас интересует не относительная эффективность нескольких видов экспериментального воздействия, а то, приведет ли оно вообще к улучшению по сравнению с исходным уровнем. Действительно, мы не можем сначала применить воздействие, а затем оценить успешность выполнения задания при его отсутствии. Но почему не сделать все в обратном порядке — то есть сначала измерить естественный уровень успешности выполнения ребенком заданий на память, применить воздействие, а затем вновь измерить память? Этот способ иллюстрирует план «Одна-группа претест-посттест» (Campbell & Stanley, 1966). Основанием для использования этого плана служит предположение, что любое повышение успешности от претеста к посттесту отражает эффект экспери-

ментального вмешательства. Если это предположение валидно, тогда нет необходимости формировать отдельные группы испытуемых.

В ряде простых ситуаций план «Одна-группа» может отвечать целям исследования. Однако, как правило, он им не отвечает. Слабость такого плана явствует из сказанного ранее об экспериментальном контроле: он допускает смешение экспериментального воздействия с рядом других факторов, которые могут обуславливать изменения от претеста к посттесту. Возьмем в качестве примера для пояснения этого момента программы вмешательства. Представьте, что мы набрали группу неблагополучных 4-летних детей, дали им тест на готовность к школе, провели с ними программу, разработанную для развития школьных навыков, и обнаружили значительное улучшение. Свидетельство эффективности нашей программы? Необязательно. Возможно, что улучшение является следствием естественного биологического созревания в ходе взросления детей — того, что Кэмпбелл и Стэнли назвали переменной **созревания**. Возможно, что улучшение является результатом других событий в жизни детей, которые произошли за время обучения по этой программе, — того, что Кэмпбелл и Стэнли назвали переменной **истории**. Возможно, что улучшение является следствием практики, полученной в ходе претеста, — того, что Кэмпбелл и Стэнли назвали переменной **тестирования**. Или же улучшение является следствием проявляющейся при любом повторном тестировании естественной тенденции к повышению первоначально низких показателей, — того, что Кэмпбелл и Стэнли назвали переменной **регрессии**. Ни одну из этих альтернативных гипотез нельзя исключить; все их можно было бы отвергнуть, если бы включалась отдельная контрольная группа, не подвергающаяся экспериментальному воздействию.

И межсубъектный, и внутрисубъектный планы принимают разнообразные формы. Ниже рассматриваются два из наиболее значимых вариантов: план уравненных групп (вид межсубъектного исследования) и план временных серий (вид внутрисубъектного исследования).

Таблица 3.3

Относительные достоинства внутрисубъектных и межсубъектных планов

Фактор	Сравнение планов
Удобство	Во внутрисубъектном исследовании — меньше испытуемых; в межсубъектном исследовании — меньше времени на каждого испытуемого
Статистические критерии	Во внутрисубъектном исследовании — более мощные, чем в межсубъектном
Эффект повторного тестирования	Присутствует во внутрисубъектном исследовании, отсутствует в межсубъектном исследовании
Возможность систематической ошибки отбора	Присутствует в межсубъектном исследовании; отсутствует во внутрисубъектном исследовании
Анализ взаимосвязей внутри группы	Обязательно во внутрисубъектном исследовании; не возможно в межсубъектном исследовании
Процедуры, производящие стойкие изменения	Обязательно в межсубъектном исследовании; невозможно во внутрисубъектном исследовании

Планы уравненных групп

Для четкого сравнения разных экспериментальных условий необходимо, чтобы испытуемые, поставленные в разные условия, были эквивалентными с самого начала исследования. Мы уже рассматривали два метода обеспечения эквивалентности: случайное приписывание разных испытуемых к разным экспериментальным группам и тестирование каждого испытуемого во всех экспериментальных условиях. Добавим теперь третий вариант: использование **планов уравненных групп**, в которых параметры испытуемых приводятся в соответствие еще до приписывания к разным экспериментальным группам.

Как мы узнали из главы 2, практически любое межсубъектное исследование предполагает определенную степень уравнивания таких явных переменных, как возраст и пол. Тогда встал вопрос: зачем ограничиваться лишь явными переменными, почему не пойти дальше и не уравнивать все потенциально значимые переменные? Небольшое замечание подскажет ответ на этот вопрос: невозможно выявить все потенциально значимые переменные, и даже если бы это было возможно, мы никогда не смогли бы добыть необходимые данные и провести необходимое уравнивание. Но все же неполное уравнивание предположительно лучше, чем никакое; почему же его не использовать? Оказывается, что эта процедура имеет как преимущества, так и недостатки.

Поскольку чаще всего при изучении детей производится уравнивание по *IQ*, я возьму эту характеристику в качестве примера. Желая уравнивать детей по параметру *IQ*, мы должны сначала провести тесты интеллекта со всеми потенциальными испытуемыми (или, возможно, обратиться к школьной картотеке, где хранятся уже собранные данные об *IQ* учащихся). Затем мы отбираем детей с одинаковыми или почти одинаковыми *IQ*. Количество детей в группах будет зависеть от количества экспериментальных условий — пары, если два типа условий, тройки — если три типа, и т. д. Работая с этими группами детей, имеющих одинаковые *IQ*, мы произвольно приписываем детей к разным экспериментальным ситуациям. Заметьте, что случайность приписывания сохраняет свое значение даже в плане уравненных групп. Заметьте также, что первоначальный подбор по *IQ* гарантирует то, чего не может гарантировать рандомизация: что в окончательном варианте экспериментальные группы будут равноценны по *IQ*.

Большим плюсом уравнивания является то, что оно обеспечивает четкий контроль переменных, которые в противном случае привели бы к появлению систематических ошибок. Если *IQ* действительно связан с величиной нашей зависимой переменной, то совершенно необходимо предотвращение смещения *IQ* и экспериментальных условий. Уравнивание также имеет ряд статистических преимуществ. Так же как и внутрисубъектные планы, план уравненных групп снижает нежелательную дисперсию и таким образом повышает мощность статистических критериев.

Основные недостатки уравнивания так или иначе касаются следующего вопроса: стоит ли оно того? Отбор, как правило, требует от исследователя приложения больших усилий, особенно если он должен предварительно протестировать всех потенциальных испытуемых (при отсутствии возможности воспользоваться уже собранными данными). Если переменная, по которой проводилось уравнивание, в

действительности не связана с величиной зависимой переменной, тогда уравнивание нам ничего не дает. Если выборка достаточно велика и используется случайное приписывание, группы, по всей видимости, при любых условиях будут эквивалентными, поэтому вновь уравнивание нам ничего не дает. Суть состоит в том, чтобы оценить эффективность затрат. Я уже отмечал, что планирование любого исследования связано с отбором нескольких подходящих приемов из большого набора потенциально информативных процедур. Тратить свое ограниченное время и усилия на процедуры, не повышающие качества исследования, — просто непрофессионально.

Помимо риска пустой траты усилий уравнивание иногда создает специфические трудности. В некоторых случаях прохождение испытуемыми предварительного тестирования может повлиять на их ответы при тестировании, имеющем экспериментальное значение (то, что Кэмпбелл и Стэнли называли переменной *реактивности*). Возможно, к примеру, что некоторые дети испытывают тревогу, когда их забирают из класса для проведения теста интеллекта, и поэтому подозрительно относятся к дружелюбному экспериментатору, который приглашает их «пойти поиграть». Попытка создать игровую атмосферу для проведения исследования поэтому может свестись к нулю, что отразится на валидности. Уравнивание иногда приводит к выбыванию испытуемых. Если испытуемых подбирать описанным выше способом, единицей измерения становится уравненная группа, а не отдельный ребенок — к примеру, тройки подобранных по *IQ* детей в исследовании с тремя экспериментальными условиями. Если кто-то из тройки по той или иной причине выбывает, то нужно исключать и двух оставшихся. Везде, где велика вероятность выпадений, произведение уравнивания может оказаться невыгодным.

Есть ситуация, в которой уравнивание — заманчивый, но обычно ложный путь. Это случай, когда исследователь хочет уравнивать изначально неэквивалентные группы. Мы уже встречались с примером этого, рассматривая различия в уровне образования между выборками молодых и пожилых людей. Обсудим еще один пример, взятый из работы исследователей Нила и Либерта (Neale & Liebert, 1986). Представьте, что вы хотите определить, добиваются ли те, кто окончил среднюю школу, большего материального благополучия, чем те, кто ее не закончил. Однако вас беспокоит, что эти две группы различаются по среднему *IQ* — допустим, 105 у выпускников и 90 у не окончивших. По причине несоответствия *IQ* любое из обнаруженных межгрупповых различий получает альтернативное объяснение: возможно, разница в благосостоянии является простым отражением различий в когнитивных способностях и не имеет ничего общего с окончанием или неокончанием школы. Поэтому вы решаете уравнивать группы по параметру *IQ*. Исключив *IQ* как потенциальную причину смещения, можно с большим основанием отнести различия в уровне благосостояния на счет преимуществ среднего образования.

Эта процедура имеет по крайней мере три недостатка, два из которых мы обсудим здесь, а рассмотрение третьего отложим до следующих разделов. Во-первых, данная процедура накладывает определенные ограничения на внешнюю валидность, поскольку по меньшей мере одна из двух групп не будет абсолютно репрезентативна (то есть либо нетипично высокий *IQ* у не окончивших школу, либо нетипично низкий *IQ* у окончивших школу). Во-вторых, приводя в соответствие

группы по одному параметру, мы тем самым систематически нарушаем соответствие по другим параметрам, связанным с окончанием школы. Предположим, вы решаете установить среднее значение *IQ* для обеих групп равным 90. В этом случае вы получите типичную группу не окончивших школу, однако ваши выпускники — в частности, потому что они преуспели несмотря на посредственный интеллект — скорее всего будут превосходить средний уровень по другим характеристикам (например, по мотивации, по поддержке со стороны семьи), которые вносят вклад в школьную успеваемость. И наоборот, установив средний *IQ* равный 105, вы получите типичных выпускников; однако теперь не окончившие школу будут иметь более низкие показатели по другим детерминантам школьной успеваемости. Приведение групп в соответствие по одному параметру может иметь незапланированный эффект, выражающийся в том, что в целом группы станут не более, а менее сходными.

Третий недостаток уравнивания неэквивалентных групп заключается в возможности эффекта статистической регрессии. В главе 4 мы обсудим то, как этот эффект проявляется при уравнивании в контексте общих рассуждений о статистической регрессии как одной из угроз для валидности.

Планы временных серий

Суть плана временных серий легче всего объяснить на примере. Целью проекта Холла и др. (Hall et al., 1971) было отучить на специальных занятиях 10-летнего мальчика громко разговаривать на уроках. Их исследование, как и все исследования по плану временных серий, состояло из нескольких фаз. Первая фаза — исходный уровень — измерение исходной частоты изучаемых поведенческих проявлений в нормальных классных условиях. Как видно из рис. 3.4, случаи такого поведения были действительно частыми — 3–4 инцидента на каждые из пяти 15-минутных уроков. За измерением исходного уровня следовало первое экспериментальное воздействие: учитель игнорировал громкие разговоры, но проявлял повышенное внимание к продуктивному поведению ребенка. Очевидным результатом введения режима «избирательного внимания» стало резкое сокращение случаев громких разговоров, что иллюстрируют результаты на втором интервале на рисунке 3.4. За фазой экспериментального воздействия следовало возвращение к исходным условиям, при которых вновь резко возросла частота нежелательных действий. Наконец, на четвертой и последней фазе было восстановлено воздействие «избирательным вниманием», и уровень нежелательного поведения вновь снизился.

Исследование Холл и др. являет собой пример плана временных серий «А–В–А–В»: фаза исходного уровня (А1), первое применение экспериментального воздействия (В1), второй исходный уровень (А2) и второе экспериментальное воздействие (В2). Проанализируем суть каждой из этих фаз. Оценка начального уровня, очевидно, необходима для выявления возможного эффекта воздействия. Фаза первого воздействия тоже, конечно, неизбежна. Но почему не остановиться сразу после того, как экспериментальное вмешательство возымело действие — то есть зачем выходить за рамки плана А–В? Ответ заключается в том, что простой план А–В не устраняет угрозы для валидности (созревание, историю и т. д.), обсуждавшиеся

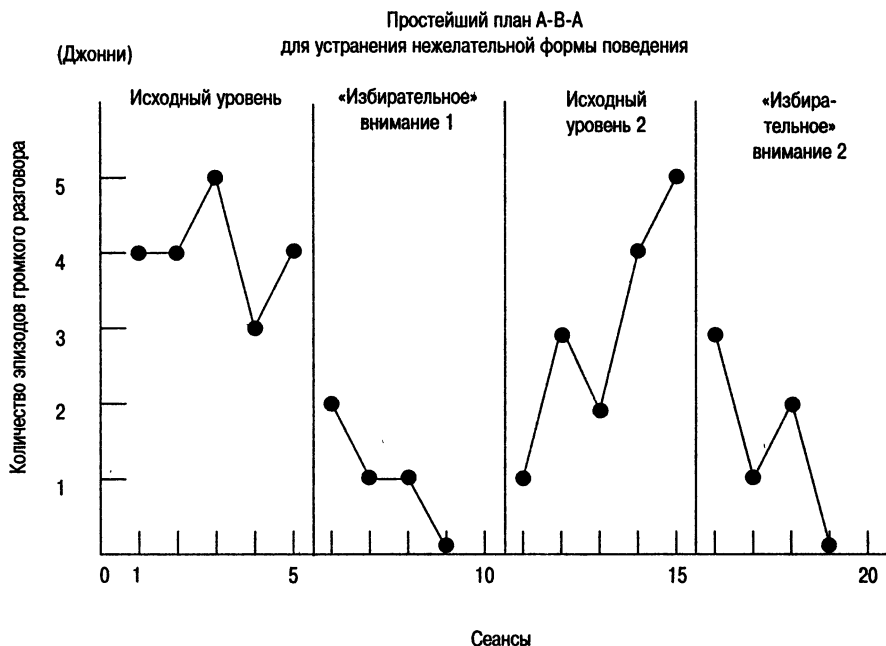


Рис. 3.4. Пример плана временных серий. Уровень корректируемого поведения (громкие разговоры) изменяется как функция от наличия или отсутствия экспериментального воздействия.
(R. V. Hall, R. Fox, D. Willard, L. *Analysis*, 4, p. 143)

ранее при рассмотрении внутрисубъективных планов в целом. В этом случае, когда у нас только один испытуемый, эти угрозы особенно трудно исключить, равно как и вероятность того, что изменение — это простое естественное колебание, не зависящее от экспериментального воздействия. Проявив, что изучаемое поведение появляется вновь, как только отменяется воздействие, мы сможем более уверенно говорить об экспериментальном воздействии как о факторе, снижающем частоту нежелательных поведенческих проявлений. Проявив, что второй период воздействия связан со вторым периодом снижения, мы удостоверимся, что причиной является экспериментальное воздействие. И разумеется, введение последней фазы В обусловлено прагматическими и этическими соображениями.

Из данного выше описания должно быть ясно, что **план временных серий** — это особая форма внутрисубъектного исследования. Он внутрисубъектный в том смысле, что каждый испытуемый проходит через все уровни независимой переменной и сравниваются результаты одного, а не разных испытуемых. Однако исследование по плану временных серий имеет и ряд отличий от внутрисубъектных исследований, описанных ранее. В большинстве внутрисубъектных исследований уровни независимой переменной представляют из себя разные формы заданий или воздействий (например, легкий и трудный материал в исследовании Дюфресна и Кобасигавы); в исследовании по плану временных серий уровни — это наличие или отсутствие экспериментального воздействия. В большинстве внутрисубъектных

исследований сравнение происходит в рамках одного экспериментального сеанса; в исследовании по плану временных серий сравнительный анализ производится в рамках повторных сеансов. В ходе большинства внутрисубъектных исследований проводится отбор и изучение групп испытуемых; многие исследования по плану временных серий (как и исследование Холла и др., 1971) направлены на изучение лишь одного испытуемого. Фактически временные серии — это главный метод, используемый в *односубъектных исследованиях* — то есть исследованиях, цель которых в выявлении эффекта экспериментального воздействия у одного испытуемого. Наконец, исследования по плану временных серий нередко проводятся по прагматическим соображениям, с целью продемонстрировать эффективность некоего способа вмешательства для коррекции определенной формы поведенческого нарушения (как в исследовании Холла и др.). Поэтому чаще всего их можно наблюдать в условиях клиники или учебных заведений.

Планы временных серий могут повлечь за собой ряд затруднений и в осуществлении, и в интерпретации, что является темой для отдельного разговора. Кроме того, они принимают множество форм, а не только описанный здесь вариант А—В—А—В. Детальное рассмотрение плана временных серий можно найти в работах Барлоу и Херсена (Barlow & Hersen, 1984), Кука и Кемпбелла (Cook & Campbell, 1979), а также Каздина (Kazdin, 1992).

Корреляционное исследование

В главе 1 я упомянул несколько актуальных, социально значимых вопросов, для решения которых оказались бы ценными исследования в области психологии развития. Приведем в качестве примера корреляционные исследования одного из таких вопросов. Маклеод, Эткин и Чафи (McLeod, Atkin & Chaffe, 1972) изучали характер воздействия насилия на телеэкране на уровень агрессии у детей. Они измерили ряд показателей агрессии у выборки 6–10-классников, а также оценили (использував самоотчеты) время, в течение которого обычно каждый ребенок смотрел по телевизору сцены насилия. Их интересовал вопрос, есть ли связь между просмотром телепрограмм с эпизодами насилия и уровнем агрессии — то есть были ли наиболее агрессивными те дети, которые наблюдали на телеэкране больше всего насилия? В их исследовании такая связь была обнаружена, что согласуется с гипотезой о провоцировании агрессии телепрограммами с эпизодами насилия.

Исследование Маклеода и др. — пример **корреляционного исследования**. Оно называется корреляционным, поскольку независимая переменная не подвергалась манипуляциям. Маклеод и др. не осуществляли экспериментальный контроль типа телепрограмм, которые смотрела их выборка, не контролировали они и уровень агрессии, которую демонстрировали дети. И параметры просмотров телепрограмм, и уровень агрессии *измерялись* в своем естественном виде; при этом цель исследования состояла в выявлении возможного совместного изменения одного показателя с другим. Эта связь может быть положительной, когда повышению одного показателя сопутствует повышение другого. Это как раз результат исследования Маклеода и др. Или же связь отрицательна, когда повышению одного показателя сопутствует понижение другого.

Результаты корреляционного исследования, как правило, оцениваются с помощью коэффициентов корреляции, о чем более подробно будет рассказано в главе 7. Пока же отметим, что коэффициент корреляции — это индекс связи между двумя переменными; его значение находится в диапазоне от -1 (абсолютно отрицательная связь) до 0 (отсутствие связи) и от 0 до $+1$ (абсолютно положительная связь). В исследовании Маклеода и др. величина коэффициента коррекции зависела от возраста и пола испытуемых, а также от использовавшихся методов оценки уровня агрессии; однако большинство показателей приходилось на интервале от $0,2$ до $0,3$, что говорит о наличии умеренно положительной связи между просмотром телепрограмм с эпизодами насилия и агрессивным поведением.

Хотя коэффициент корреляции обычно связан с корреляционными планами, важно разграничивать статистический показатель и план. Помимо коэффициентов корреляции для анализа результатов корреляционного исследования используются и другие статистические показатели. Маклеод и др., к примеру, могли бы разделить свою выборку на группы тех, кто смотрит телевизор часто, умеренно часто и редко, а затем использовать для сравнения уровня агрессии в трех группах t -критерии или дисперсионный анализ. В этом случае статистические показатели были бы другими, однако план остался бы корреляционным. По причине независимости плана и статистических показателей, некоторые исследователи предпочитают называть данный вид исследования неэкспериментальным. Каков бы ни был термин, суть исследований остается в том, что переменные просто измеряются, а не контролируются экспериментальным путем.

Корреляция и причинность

Нет необходимости доказывать, что корреляция не тождественна причинности. То есть из того, что две переменные коррелируют, мы не можем сделать вывод о характере причинно-следственных связей между ними, если таковые вообще имеются. Поэтому результаты исследования Маклеода и др. согласуются с гипотезой о том, что просмотр телепрограмм с эпизодами насилия вызывает агрессию, однако не могут доказать истинность этой гипотезы.

Указанное базовое ограничение корреляционного исследования обусловлено отсутствием экспериментального контроля. Как уже неоднократно подчеркивалось, именно контроль — контроль характера независимой переменной, контроль включения испытуемых в разные экспериментальные группы, контроль потенциально значимых переменных — гарантирует внутреннюю валидность выводов о возможных причинно-следственных связях. Поскольку в корреляционном исследовании эти формы контроля отсутствуют, максимум, что оно может продемонстрировать, это совместное изменение двух или более показателей. Но оно не объяснит, почему происходит данное изменение.

Рассмотрим исследование Маклеода и др. Выявленную в нем корреляционную связь можно объяснить тремя способами. Первый: просмотр телепрограмм с эпизодами насилия является причиной повышения агрессии. Если бы исследователи экспериментально изменили параметры просмотров, то могли бы довольно уверенно сделать такой вывод. Но поскольку экспериментальные манипуляции не производились, существуют альтернативные объяснения: возможно, более агрессив-

ные дети выбирают более агрессивные телепрограммы. В этом случае уже склонность к агрессии является причиной просмотра программ с эпизодами насилия, а не наоборот. Наконец, есть и третье объяснение: возможно, и просмотр агрессивных телепередач, и агрессивное поведение определяются неким третьим причинным фактором, а не влияют друг на друга. К примеру, определенные стили воспитания способствуют и агрессивности, и предпочтению агрессивных телепрограмм; поэтому два показателя изменяются параллельно, однако ни один не является причиной другого.

Это доказательство можно изложить и в общем виде. Всегда, когда между переменной A и переменной B существует корреляционная связь, следует рассматривать три альтернативных варианта: A является причиной B , B является причиной A или некий третий фактор C является причиной A и B .

Непригодность для выявления причинно-следственных связей существенно ограничивает возможности корреляционных планов. Зачем тогда их использовать? Главное основание состоит в том, что в некоторых случаях такое исследование — это лучшее, что можно сделать. Многими переменными нельзя манипулировать из этических или практических соображений — методами воспитания, к примеру, или использованием психоактивных веществ в пренатальный период. При таких обстоятельствах единственным выходом является корреляционное исследование. В других случаях экспериментальное воздействие возможно, но затруднительно, особенно, если цель — сочетание экспериментального контроля и естественной обстановки. Эту мысль иллюстрирует пример с насилием на телеэкране и агрессией. Экспериментально изменять параметры просмотра, а затем измерять параметры агрессии *возможно*, и во многих исследованиях это проделано, однако все эти исследования заслуживают критики из-за искусственности условий и низкого уровня внешней валидности. В исследованиях же, подобных описанному выше, в центре внимания находятся непосредственно две изучаемые переменные: естественный просмотр телепрограмм и естественная агрессия. Еще одним достоинством корреляционного подхода является то, что он позволяет учесть более широкий диапазон изменений. В экспериментальном исследовании просмотра телепрограмм и агрессии мы были бы вынуждены ограничиться созданием двух-трех типов экспериментальных условий. Корреляционное же исследование охватывает весь диапазон естественного опыта, от 2–3 часов в неделю на одном конце континуума до 40–50 часов в неделю на другом.

Способы повышения обоснованности выводов о причинно-следственных связях

По результатам корреляционного исследования нельзя с уверенностью говорить о характере причинно-следственных связей. Однако есть методы повышения обоснованности выводов о любых причинных воздействиях. В этом разделе мы рассмотрим некоторые из подобных методов.

Первая стратегия основана на простом здравом смысле, однако стоит того, чтобы о ней упомянуть. В некоторых случаях одно из направлений причинного воздействия исключается самим характером переменных. Предположим, мы обнаружили положительную корреляционную связь между ростом ребенка и уровнем

агрессии. Вполне допустимо, что рост ребенка каким-то образом влияет на уровень агрессии (хотя нам все еще необходимо было бы уточнить, как именно). Однако вряд ли уровень агрессии оказывает какое-либо причинное воздействие на рост ребенка. В таких случаях нужно принимать во внимание всего две гипотезы: *A* является причиной *B* или *C* является причиной *A* и *B*. Возможность влияния *B* на *A* не учитывается.

Подход с точки зрения логики используется в вопросах, касающихся направления причинно-следственных связей между *A* и *B*. Второй метод особенно эффективен для исключения третьего фактора *C*. При этом пользуются аналитической процедурой, которая называется **методом парциальной корреляции**. Частичная корреляция — это процедура, направленная на статистическое устранение действия одной переменной на корреляционную связь между двумя другими переменными. По существу, метод частичной корреляции заключается в сохранении неизменной потенциально значимой третьей переменной в процессе анализа связи между двумя другими переменными. Это равноценно изучению отношений между *A* и *B* в выборке, где все имеют одинаковые показатели по параметру *C*. Вопрос, разумеется, состоит в том, остается ли значимой связь между *A* и *B* при отсутствии влияния *C*.

Предположим, мы обнаружили положительную корреляционную связь между просмотром телепрограмм с эпизодами насилия и уровнем агрессии, однако подозреваем, что эту связь в действительности обуславливает некий третий фактор, к примеру методы воспитания. Допустив, что мы можем получить приемлемую оценку способов воспитания, мы могли бы использовать метод частичной корреляции для устранения влияния воспитания на корреляционную связь «просмотр телепрограмм — агрессия». Обнаружив, что корреляция остается на том же или примерно на том же уровне, какой она была изначально, мы могли бы заключить, что способ воспитания не является значимым фактором, привносящим смещение. И наоборот, резкое снижение корреляции свидетельствовало бы о наличии существенной доли воспитания в связи между просмотром агрессивных телепрограмм и агрессивным поведением.

Хотя процедуры и различаются, цель использования метода парциальной корреляции та же, что и цель метода уравнивания, описанного ранее в этой главе. В обоих случаях исследователь стремится устранить факторы, вносящие смещение, приводя их к одному уровню для всех сравниваемых групп. При уравнивании приведение к одному уровню осуществляется до проведения процедуры исследования путем включения испытуемых в разные группы; при парциальной корреляции оно осуществляется после процедуры исследования путем статистического исключения факторов, приводящих к смещению. И уравнивание, и частичная корреляция имеют общее ограничение: с их помощью нельзя устранить *все* возможные факторы, приводящие к смещению. Иными словами, существует масса переменных *C*, и ни одному исследователю не измерить и не проконтролировать их все.

Процедуры третьего типа, предназначенные для вычленения информации о причинно-следственных связях из корреляционных данных, так или иначе связаны с временными отношениями между переменными. Эти процедуры строятся на факте необходимого предшествования причины следствию. Поэтому, составив

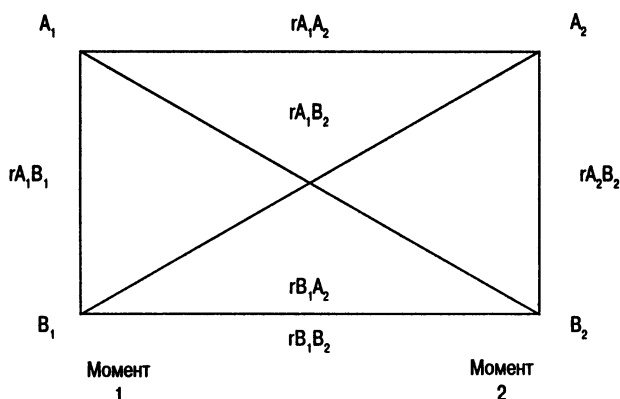


Рис. 3.5. Корреляции между переменными в плане перекрестных корреляций с лагом

схему изменений отношений A и B во времени, мы сможем ближе подойти к решению вопроса о том, A приводит к B или наоборот.

Наиболее трудоемкой формой временного анализа является процедура, называемая **перекрестной корреляционной группировкой с лагом** (Campbell & Stanley, 1966). Она требует лонгитюдного подхода, при котором по крайней мере две переменные измеряются два или более раз за определенный период времени. Такое исследование дает возможность вычислить ряд коэффициентов корреляции как в рамках одного периода, так и между периодами. На рис. 3.5 представлены варианты корреляций в исследованиях с двумя переменными и двумя моментами измерения (r -коэффициенты корреляции). Таким образом, можно узнать коэффициент корреляции между A и B в момент 1, и в момент 2; стабильность A и B во времени; и, что наиболее важно, межвременные коэффициенты корреляции между A и B — то есть корреляции, изображенные по диагоналям. Для определения направления причинного воздействия в паре A и B необходима информация именно по диагонали. Если причинный фактор это A , то следует ожидать значительной корреляции между A в момент 1 и B в момент 2 — изменения в причине должны приводить к изменению эффекта. Коэффициент корреляции между B в момент 1 и A в момент 2 должен быть существенно ниже. И наоборот, если причинный фактор — B , следует ожидать более высокого коэффициента корреляции между $B1$ и $A2$, и более низкого — между $A1$ и $B2$.

Наверное, проще всего разъяснить вышесказанное на практическом примере. Лефковитц, Эрон, Уолдер и Хьюсман (Lefkowitz, Eron, Walder & Huesmann, 1972) использовали метод перекрестных корреляций, изучая связь насилия на телеэкране и агрессии (см. также Huesmann, Lagerspetz & Eron, 1984). Они измерили параметры просмотров телепрограмм и оценили уровень агрессии у 3-классников, а затем у них же (у тех, кого удалось привлечь к повторному обследованию) и те же параметры 10 лет спустя. Результаты для мужской половины выборки представлены на рис. 3.6. Обратите внимание: ничто не подтверждает гипотезу о том, что склонность к агрессии ведет к предпочтению телепрограмм с эпизодами насилия;

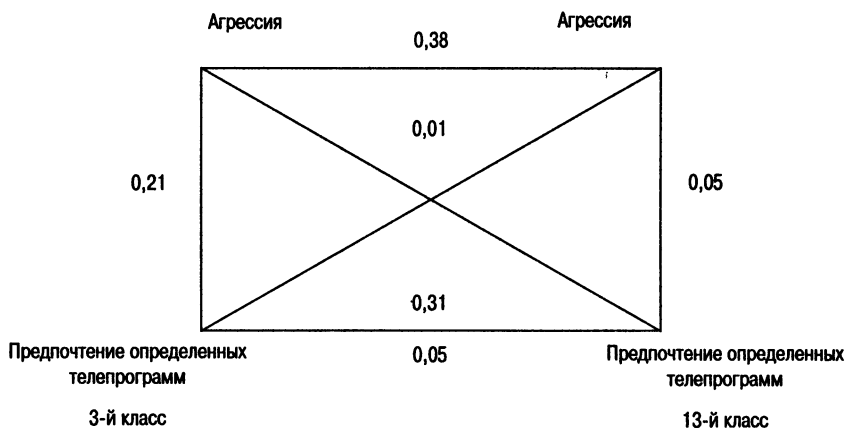


Рис. 3.6. Корреляции между предпочтением определенных телепрограмм и агрессией в перекрестном корреляционном исследовании с лагом.
(M. M. Lefkowitz, L. D. Eron, L. D. Walder, & L. R. Huesmann, 1972, *Television and Social Behavior*. Vol. 3)

коэффициент корреляции между уровнем агрессии в момент 1 и просмотром агрессивных телепередач в момент 2 по существу нулевой. Однако есть свидетельства в пользу противоположной гипотезы, поскольку просмотр агрессивных телепередач в момент 1 тесно связан с агрессией в момент 2.

Следует отметить, что исследование Лефковитца и др. было описано исключительно в качестве примера перекрестной группировки с лагом. В действительности это исследование и сделанные на его основе выводы можно подвергнуть критике по ряду пунктов (Freedman, 1984). В целом, перекрестное исследование с лагом, хотя и выглядит довольно простым в том виде, в котором оно здесь рассмотрено, связано с рядом сложностей и зависит от некоторых статистических допущений. Неизвестно, как часто эти допущения соответствуют действительности и, следовательно, насколько обосновано применение данного метода. Рогоса (Rogosa, 1980) аргументированно критикует попытку сделать выводы о причинности посредством перекрестного анализа.

Рассмотрение последнего метода повышения обоснованности заключений о причинно-следственных связях поможет уяснить один из базовых принципов методологии исследования. Иногда есть возможность дополнять корреляционный анализ экспериментальным изучением того же вопроса. Иными словами, можно изменить предположительную причину и оценить воздействие данных манипуляций на зависимую переменную, таким образом создавая реальную связь между зависимой и независимой переменными. Как уже отмечалось, в литературе, посвященной теме насилия на телеэкране, содержится описание ряда подобных экспериментальных исследований, в которых оценивается эффект манипуляций со временем и характером телепрограмм. В этих исследованиях осуществляется контроль, отсутствие которого является недостатком корреляционного плана. Наличие переменной *A*, подвергающейся экспериментальным манипуляциям, снимает не-

определенность в отношении направления причинного воздействия в паре *А* и *В* — изменения *В* должны быть обусловлены изменениями *А*, а не наоборот. А поскольку мы контролируем не только нашу независимую переменную, но и другие факторы, никакая третья переменная *С* не может вносить смещение в отношения между *А* и *В*. Поэтому наши выводы о причинно-следственных связях будут более обоснованы.

Принцип, который иллюстрирует этот пример, касается значимости конвергентных операций при исследовании комплексных, сложных для изучения объектов. Термин **конвергентные операции** (другое название — *мультиметодический подход* — например, Brewer & Hunter, 1989) означает использование, либо в рамках одного, либо в рамках ряда исследований, разнообразных методов изучения определенного вопроса (игнорирование этого принципа — использование только одного метода ставит под угрозу валидность, что Кук и Кемпбелл, 1979, обозначили как **монометодическое искажение**). Основная мысль заключается в том, что слабые стороны одного метода могут, по крайней мере отчасти, компенсироваться сильными сторонами другого и что выводы, основанные на совмещении результатов, полученных с помощью разных методов, более достоверны, чем выводы, полученные при помощи только одного метода. Этот довод, несомненно, касается и вопроса о взаимосвязи насилия на телеэкране и агрессии. Здесь для выявления причинно-следственных связей единственно верным способом является эксперимент; в то же время экспериментальное исследование страдает рядом недостатков (искусственность, эффект реактивности и т. д.), которые ставят под сомнение его внешнюю валидность. В корреляционных планах удастся избежать множества погрешностей исследований с использованием манипуляций; однако, как мы увидели, корреляционное исследование по существу не может дать информации об отношениях между причиной и следствием. Именно в связи с этими ограничениями каждого из методов в отдельности необходимо совмещение результатов использования разнообразных методов. Поэтому корреляционные исследования влияния телевидения дают нам возможность с большей уверенностью говорить о том, что влияние сцен насилия на телеэкране, продемонстрированное в экспериментальных условиях, существует и в реальной жизни. Аналогично тот факт, что экспериментальное изменение характера просмотра телепрограмм отражается на уровне агрессии детей, дает нам право утверждать, что телевидение действительно является причиной в корреляционной паре телевидение—агрессия.

Резюме

В этой главе рассматриваются три темы, попадающие в разряд исследовательского плана: сравнение возрастных групп, сравнение экспериментальных условий и различие между экспериментальными и корреляционными планами с особым акцентом на вопросе корреляции и причинности.

При изучении разных возрастных групп чаще всего используются два плана: лонгитюдный план и план поперечных срезов. В *лонгитюдном исследовании* на протяжении определенного периода времени регулярно обследуются одни и те же

испытуемые. Только этот подход дает возможность измерить непосредственно возрастные изменения, а не возрастные различия. Это также единственный способ изучения индивидуальной стабильности или изменчивости во времени. Но с другой стороны, лонгитюдное исследование требует больших затрат времени и усилий, что, несомненно, объясняет относительно редкое его использование. Лонгитюдное исследование также приводит к ряду систематических ошибок. К ним относятся *избирательный отсев* испытуемых в ходе исследования, эффект *тестирования*, обусловленный повторным тестированием с использованием одних и тех же методик, и неизбежное смещение возраста испытуемого и времени тестирования.

В исследовании методом *поперечных срезов* изучаются разные испытуемые разных возрастов. Этот подход, как правило, экономичнее лонгитюдного; с его помощью можно избежать множества проблем лонгитюдного исследования, он вполне пригоден для изучения ряда тем. Однако поперечные планы имеют и недостатки. Поскольку все испытуемые обследуются только по одному разу, невозможно оценить непосредственно возрастные изменения. Возрастное сравнение также затрудняет *систематическая ошибка отбора* при формировании возрастных групп, а также *избирательный отсев* испытуемых из исследования. Еще одна проблема, характерная как для лонгитюдных, так и для поперечных планов, — проблема *эквивалентности измерений* отбор методик, равноценных для сравниваемых возрастных групп. Наконец, в поперечных планах неизменно возникает смещение между возрастом испытуемых и поколением, или когортой, к которым они принадлежат.

Недостатки классических лонгитюдного и поперечного методов обусловили разработку альтернативных планов. В *плане с временным лагом* при изменении параметров поколения и времени измерения неизменным сохраняется возраст испытуемых. При помощи такого плана можно оценить значение факторов, которые в традиционных планах смешиваются с возрастом. Более трудоемки *последовательные планы*, сочетающие элементы лонгитюдного исследования, исследования методом поперечных срезов и плана с временным лагом. Последовательные планы, несомненно, информативнее; они, однако, требуют и больших затрат, при этом не устраняя все возможные источники смещения.

Второй раздел главы посвящен планам сравнений разных заданий или экспериментальных условий. Существует два основных подхода: *внутрисубъектный* в котором каждый испытуемый выполняет все задания (или находится во всех типах экспериментальных условий); и *межсубъектный*, при котором разные испытуемые должны выполнять разные задания (или находиться в разных условиях). Внутрисубъектное исследование, как правило, более экономично, статистически более достоверно и позволяет избежать ряда затруднений (например, систематической ошибки отбора), с которыми может столкнуться межсубъектное исследование. Внутрисубъектный подход незаменим для изучения внутригрупповых взаимосвязей и закономерностей. Межсубъектный подход, в свою очередь, позволяет избежать ряда проблем внутрисубъектного исследования — в частности, *эффектов последовательности и переноса*, обусловленных повторным тестированием. Межсубъектный подход незаменим в том случае, когда экспериментальное воздействие призвано произвести некоторые стойкие изменения.

Затем были рассмотрены конкретные варианты межсубъектных и внутрисубъектных исследований. В *плане уравненных групп* параметры испытуемых приводятся в соответствие до приписывания их к разным экспериментальным условиям. Преимуществом уравнивания является то, что оно обеспечивает эквивалентность групп по переменным (например, *IQ*), которые могли бы повлиять на результаты. К недостаткам этого метода относятся увеличение затрат времени и усилий, искажающее влияние предварительного тестирования, более высокий уровень выбывания испытуемых, обусловленный необходимостью исключать всех подобранных по одному признаку при выбывании любого из них, а также вероятность того, что приведение в соответствие по одному параметру будет систематически нарушать соответствие между другими переменными. В *плане временных серий* экспериментальное воздействие применяется периодически, при этом регистрируются изменения в поведении как функция от наличия или отсутствия воздействия. Такой вид исследования чаще всего проводится в клинике или в условиях учебных заведений, зачастую в форме исследования с одним испытуемым.

Главу завершает рассмотрение *корреляционного исследования*. В корреляционном плане независимые переменные не контролируются; измеряются две или более переменные и оценивается степень их совместной изменчивости. Корреляционный план — единственно приемлемая форма исследования, если изучаются переменные, манипулирование которыми невозможно или затруднительно. Кроме того, корреляционное исследование способно охватить более широкий диапазон уровней переменной. С другой стороны, отсутствие экспериментального контроля подразумевает невозможность корреляционного исследования дать информацию о причинно-следственных связях. К методам, используемым для снижения неопределенности и для уточнения характера причинно-следственных связей, относят логический анализ возможных направлений причинности; *частичную корреляцию*, при которой статистически устраняется влияние третьих факторов; *перекрестную корреляционную группировку с лагом*, в которой оцениваются корреляционные связи во временной перспективе; и экспериментальные манипуляции с одной из переменных.

Упражнения

Одна из тем этой главы — трудность разграничения эффектов возраста и эффектов когорты, или поколения. Возьмите в качестве примера когорту, к которой вы принадлежите. Есть ли что-нибудь в вашем жизненном опыте, что отличало бы его от опыта представителей других поколений? Какого рода эффект могло бы иметь это различие при сравнении методом поперечных срезов?

Один из способов анализа последовательных планов — представить отдельные результаты и подумать, что они могли бы значить. Проанализируйте изображенный ниже план последовательности когорт (рис. 3.7). Зависимая переменная — это *IQ*; допустим, что средние значения *IQ* для разных групп колеблются в пределах от 90 до 110. Для каждого из нижеследующих результатов продумайте способы выявления: а) эффектов только возраста; б) эффектов только когорты; в) эффектов возраста и времени измерения.

		Время измерения		
Когорты		1990	2000	2010
		Воз- раст	Воз- раст	Воз- раст
	1930	60	70	80
	1940	50	60	70
	1950	40	50	60

Рис. 3.7. План последовательности когорт

В этой главе подчеркнуты как достоинства, так и недостатки лонгитюдного исследования. Одним из альтернативных подходов к изучению стабильности/изменчивости во времени является *ретроспективный метод*. Ретроспективный метод предполагает анализ прошлого, обычно взрослых людей, обладающих интересующими исследователя качествами, попытку выявить существенные предпосылки или определяющие факторы раннего развития, которые обусловили имеющийся результат. Очевидная трудность в этой ситуации — получение точных данных о прошедших событиях. Выберите какой-либо значимый опыт из своего прошлого — успехи в школе, к примеру, или отношения с друзьями, или путешествия вместе с семьей. Восстановите по возможности наиболее точно картину развития на определенном отрезке периода детства. Попросите сделать то же самое своих родителей и сравните ваши воспоминания. Если есть возможность, сравните оба рассказа с объективными свидетельствами (например, табелем успеваемости, семейным фотоальбомом).

Как отмечалось в тексте, корреляция между переменными *A* и *B* может интерпретироваться несколькими способами: *A* является причиной *B*, *B* является причиной *A*, некоторый третий фактор *C* является причиной *A* и *B*, или сочетание всех этих вариантов. Ниже представлено несколько примеров положительных корреляций, полученных в реальных исследованиях. Для каждого: а) предложите как можно больше правдоподобных объяснений; б) перечислите, какие данные помогли бы вам выбрать из возможных объяснений:

- использование родителями физического наказания и агрессия ребенка;
- разъяснение родителями норм поведения и просоциальное поведение ребенка;
- физическая привлекательность и популярность;
- *IQ* и школьная успеваемость;
- представление о себе как об учащемся и школьная успеваемость;
- уровень активности и умственная сохранность в пожилом возрасте.

Измерение

В главе 2 было рассмотрено кардинальное различие между независимой и зависимой переменными: независимая переменная — это те факторы, которые мы контролируем; зависимая переменная — это то, что мы измеряем на выходе. Большая часть главы 3 «План» была посвящена первому из этих понятий — разнообразным способам создания независимых переменных и их сочетанию. В этой главе акцент будет смещен с независимой на зависимую переменную: на способы измерения результатов исследовательского процесса.

Измерение — это большая тема; в ходе изложения материала мы будем периодически к ней обращаться. В частности, главы, посвященные специфическим видам исследования (главы 10–13), содержат достаточный объем информации о том, как измерять конструкторы, наиболее интересные для специалистов в области психологии развития. Цель этой главы — лишь познакомить с некоторыми базовыми принципами измерения — принципами, к которым мы по необходимости будем возвращаться и которые будем конкретизировать.

Структура этой главы следующая. В первом разделе вводится ряд базовых понятий, необходимых для понимания измерения, к примеру, понятие операционального определения, а также такие основополагающие понятия, как надежность и валидность измерения. Оставшаяся часть главы посвящена двум важнейшим формам измерения: стандартизованным тестам для оценки определенных психологических качеств и методам наблюдательной оценки поведения. Рассматривая эти типы измерения, мы обсудим еще ряд общих принципов.

Базовые понятия

Как и ранее, многие принципы будут рассмотрены на конкретном примере, что облегчит их понимание. Он относится к области, о которой уже не раз говорилось: насилие на телеэкране и агрессия. Либерт и Бэрн (Liebert & Baron, 1972) использовали для изучения этого вопроса экспериментальный план, показывая половине своих испытуемых (5–9-летним детям) сцены из агрессивного фильма (3,5-минутный ролик из «Неприкасаемых»), а другой половине — ролик нейтрального содержания, после чего сравнивали уровень агрессии в обеих группах. Агрессия

оценивалась по множеству критериев, но один из них служил основой для большинства выводов; на нем мы и сосредоточимся. После просмотра детей сажали перед нультом с двумя кнопками: красная кнопка означала «боль», а зеленая — «помощь». От нуля через отверстие в стене шли провода, и детям сообщали, что пульт подключен к игровому автомату, с которым в соседней комнате будет играть другой ребенок. По правилам игры нужно было поворачивать рычаг, а связь пульта с автоматом позволяла испытуемым воздействовать на ход игры: нажатие на кнопку «помощь» облегчало поворот рычага, а нажатие на кнопку «боль» делало рычаг горячим, таким образом затрудняя игру. Затем следовала серия попыток, когда ребенок мог нажимать на любую из кнопок столько, сколько захочет. Предстояло выяснить, будут ли дети, которые смотрели ролик с эпизодами насилия, проявлять большую агрессию, дольше и чаще нажимая на кнопку «боль». Именно это и произошло: длительность нажатия кнопки «боль» у детей из группы, смотревших агрессивный ролик, была почти на 50 % больше, чем у детей, смотревших ролик нейтрального содержания. Таким образом, результаты подтвердили предположение о том, что просмотр детьми агрессивных телепрограмм ведет к повышению агрессии. (Добавим лишь то, о чем вы, вероятно, уже догадались: в соседней комнате не было никакого ребенка.)

Операционализация

Рассмотрим два способа подведения итогов по результатам исследования Либерт и Бэрона: «Просмотр детьми телепрограмм с эпизодами насилия повышает их агрессивность» и «Просмотр детьми 3,5-минутного ролика, составленного из отрывков фильма “Неприкасаемые” в лабораторных условиях повышает вероятность того, что пять минут спустя дети будут нажимать кнопку, нажатие на которую будто бы приводит к болезненным ощущениям ребенка, находящегося в соседней комнате и не видимого для испытуемого». Очевидно, что первое утверждение представляется более интересным и обобщенным. Однако второе — более точно, так как описывает именно то, что было проделано и обнаружено, тогда как первое утверждение — это общий вывод, выходящий за рамки фактических данных.

Два указанных выше способа подведения итогов в исследовании Либерт и Бэрона демонстрируют существенное различие между выводами, которые хочет сделать исследователь, и конкретными манипуляциями и измерениями, которые он проделывает. «Насилие на телеэкране» и «агрессия», несомненно, интересные конструкты, стоящие того, чтобы их изучать. Таковы и многие другие вещи, которые изучают специалисты в области психологии развития — интеллект, креативность, Я-концепция, половая типизация. Проблема заключается в том, что интеллект и креативность, в действительности, не «вещи», произвольно и непосредственно наблюдаемые; прежде чем изучать, их нужно неким образом *операционализировать* — то есть перевести в специфическую, измеряемую форму, а любое измерение требует превращения общих понятий в специфические.

От глагола *операционализировать* образуется существительное — **операциональное определение**. Понятие операциональное определение впервые появилось в 20-х годах XX века в работах физика П. У. Бриджмена (P. W. Bridgemen). Операциональное определение — это определение переменной в терминах операций, используемых для создания или измерения этой переменной. Так, температуру можно определить как перемещение ртути внутри определенного вида емкости.

Интеллект можно определить как успешность выполнения заданий теста Стэнфорд—Бине, агрессию можно определить по параметрам, указанным во втором из итоговых заключений. Во всех случаях налицо привязка к используемым на практике измерительным операциям.

При узком толковании операционализация предполагает, что исследователь, делая выводы об изучаемых переменных, не выходит за рамки операций, использовавшихся для их создания или измерения. Фактически сегодня мало кто из исследователей придерживается столь буквального понимания операционализации. Тем не менее, операционализм оказал значительное и благотворное влияние на психологическую науку. Он установил общепринятую схему, в границах которой решается задача измерения. Основной принцип этого подхода заключается в требовании четкости определения, объективности измерительных операций и их воспроизводимости любым исследователем в любой, соответствующим образом оборудованной лаборатории.

Рассмотрим, как может происходить перевод теоретического конструкта в конкретные измеримые параметры для исследователя и для тех, кто читает исследовательские отчеты. Возьмем, к примеру, обсервативное исследование агрессии в детском саду. Исследователя интересует вероятность того, что социальное подкрепление способствует агрессивному поведению в обстановке детского сада. Первая задача исследователя состоит в выборе операционального определения каждого из этих довольно глобальных конструктов. Поскольку любой из них операционализируется множеством способов, из совокупности свойств данных конструктов нужно выбрать ограниченное число конкретных параметров. Наш исследователь, к примеру, может определить социальное подкрепление как набор неких вербализаций (например, «хорошо», «о'кей»), неких выражений лица (например, адресованная ребенку улыбка) или определенного рода невербальное поведение (например, похлопывание по плечу, объятия). Агрессию можно определить как набор разнообразных физических актов (например, удары, пинки, шипки), направленных на причинение вреда другому человеку. Каковы бы ни были выбранные параметры, задача исследователя состоит в проведении максимально точных измерений, в описании именно того, что было сделано.

Перед тем, кто в итоге будет знакомиться с исследовательским отчетом, тоже стоит определенная задача. Он должен прежде всего осознать, что такие конструкты, как социальное подкрепление и агрессия, могут иметь множество операциональных определений и что какое-то одно исследование неизбежно будет включать лишь часть возможных определений. Это означает, что используемые операциональные определения не обязательно будут согласовываться с тем, что читатель сам вкладывает в понятия социального подкрепления и агрессии, и что определения не обязательно будут согласовываться с теми определениями, которые читатель встречал в других исследовательских отчетах по этой проблематике. Таким образом, читатель должен отбросить, по крайней мере на какое-то время, составленные ранее представления и сосредоточиться на том, что действительно было осуществлено в рассматриваемом исследовании. Тогда он сможет прийти к заключению, что операциональные определения разумны и достаточны, либо что они по тем или иным причинам неадекватны. К примеру, в категорию социального подкрепления следует включить лишь заботу о ребенке, а также очевидные позитивные поведенческие проявления, такие как улыбка и похвала. В категорию агрессии следует

включить вербальные, а не только физические действия. Возможно, одно из наиболее важных качеств, которые должен развить у себя читатель психологической литературы, — это способность выйти за пределы гладких выводов из «Реферативных журналов» и «Актуальных вопросов...» (например, социальное подкрепление приводит к повышению уровня агрессивности) с тем, чтобы оценить исследование в терминах использованных в нем на практике операций. Если эти операции не отвечают определенным требованиям, общие выводы вряд ли можно считать бесспорными.

Квантование

Для измерительной системы характерен переход от глобального к специфическому. Еще одним ее свойством является *квантование*. По словам одного из создателей теории измерения С. С. Стивенса (S. S. Stevens), «измерение — это численное выражение определенных аспектов объекта в соответствии с тем или иным правилом» (Stevens, 1968, р. 850). Однако числа и правила могут носить разный характер, в зависимости от используемых форм измерения. Соответственно изменяются и выводы, которые можно сделать на основе полученных результатов.

Типы величин определяют так называемые *уровни*, или *шкалы измерения*. Вслед за Стивенсом (например, Stevens, 1968) ученые традиционно выделяют четыре уровня измерения. Каждый уровень выполняет базовые функции любой системы измерения, то есть результатом каждого измерения является некая величина, а полученные величины позволяют разграничить результаты всех измерений. Однако происходит это по-разному.

Наиболее простая из систем измерения — **номинальная шкала**. *Номинальная* это то же, что «именующая» — придающая некоторое качественное значение каждому результату измерения. Предположим, вас интересует, какие игрушки нравятся дошкольникам. Вы предлагаете каждому ребенку четыре игрушки и просите выбрать ту, с которой ему больше всего хотелось бы поиграть. Измерение состоит в регистрации того, какая из игрушек была выбрана. Измерение в этом случае *номинативное*, поскольку вы даете «имя» каждому ответу. Конечно, вы можете превратить названия в числа — к примеру, записать 1, когда выбор падает на грузовик, 2, когда оказывается предпочтение медвежонку, и т. д. Эти числа, однако, будут выполнять функцию наименований, не имея какого-либо количественного значения. Данный факт является отличительной особенностью *номинативных шкал*: они позволяют классифицировать по качеству, но не по количеству.

Рассматривая второй уровень измерения, можно возвратиться к примеру с агрессией. Предположим, мы попросили воспитателя проранжировать детей по степени агрессивности каждого из них в игровой комнате. Мы использовали рейтинговую шкалу, имеющую пять позиций: от «крайне агрессивного» до «умеренно агрессивного» и от «умеренно агрессивного» до «совершенно неагрессивного». Системой измерения служила **порядковая шкала**, поскольку мы располагали результаты наблюдений по их величине. В этом случае, в отличие от случая использования *номинативной шкалы*, результаты измерений имеют количественную размерность, а измерение служит для определения порядковой позиции каждого из результатов. Поэтому можно сказать, что «крайне агрессивный» ребенок агрессивнее «умеренно агрессивного», который, в свою очередь, агрессивнее «совершенно неагрессивного» — или что пятая категория действительно больше третьей, кото-

рая, в свою очередь, больше первой. Заметьте, однако, что все еще нельзя говорить о масштабах различий. Мы не знаем, к примеру, является ли разница между пятой и третьей позицией такой же, как разница между третьей и первой. Не вправе мы конечно и утверждать, что ребенок, попавший в пятую категорию, в пять раз агрессивнее ребенка, попавшего в первую. Все, о чем можно говорить, — это о порядке.

Данное ограничение снимается в третьей шкале, интервальной. В **интервальной шкале** точки деления не только упорядочены, но и равноудалены друг от друга. Характерный (хотя и не психологический) пример — шкала термометра. Температурные показатели, несомненно, упорядочены: 40° теплее, чем 30° , а 30° теплее, чем 20° . Более того, точки деления на термометре расставлены равномерно. Поэтому мы можем утверждать, что разница между 40° и 30° точно такая же, как и разница между 30° и 20° (в физическом, если не психологическом смысле). Как было сказано выше, такого рода количественная точность невозможна на порядковой шкале.

Однако интервальная шкала все же имеет одно ограничение: на ней нет истинного нулевого деления. На термометре, конечно, есть ноль, однако ноль на шкале термометра — это произвольная точка, с обеих сторон которой имеются другие значения, а не действительно нулевой уровень. Это не подразумевает полного отсутствия измеряемых характеристик. Измерительные шкалы, отвечающие всем критериям интервальных шкал и, кроме того, содержащие истинный ноль, называются **шкалами отношений**. Характерным примером шкал отношений могут служить системы измерения таких физических параметров, как высота или вес. Шкала весов включает не только равные интервалы между значениями веса, но и истинный ноль, то есть отсутствие веса на весах. Наличие нулевой точки на шкалах отношений позволяет делать выводы о пропорциях, что невозможно при использовании интервальных шкал. К примеру, можно сказать, что 40 фунтов в 2 раза тяжелее 20 фунтов. Но нельзя сказать, что 40° в 2 раза теплее 20° .

Уровень измерения — один из факторов, определяющих выбор статистических критериев. К вопросу об измерительных шкалах мы вернемся в главе 7, посвященной статистике.

Аспекты измерения

До сих пор мы акцентировали внимание в основном на необходимости осуществлять выбор при переводе некоего глобального конструкта (например, агрессии) в некоторую специфическую, измеряемую форму (например, удары и пинки). Однако мы еще не обсудили параметры, по которым осуществляется измерение. Теоретики выделяют ряд параметров, или «граней» измерения (Messick, 1983). Часть из них мы рассмотрим в этом разделе, остальные — в дальнейшем.

Одно из кардинальных решений, которое нужно принять исследователю, касается определения конкретного аспекта поведения, который он будет изучать. Предположим, что наш воображаемый исследователь агрессии у дошкольников выбрал в качестве критерия агрессии удары. Однако все еще остается нерешенным вопрос, что конкретно в этих ударах будет измеряться. Исследователь, к примеру, может решить работать с *частотой действий*, то есть оценивать, как часто ребенок наносит удары. Такой прямой подсчет, вероятно, даст наиболее очевидный индекс того, что принято подразумевать под «уровнем агрессии». Альтернативный вариант —

работать не с частотой, а с *интенсивностью*, то есть оценивать, не как часто ребенок наносит удары, а какой силы эти удары. Интенсивность также имеет очевидную связь с тем, что понимается под «агрессией». Есть и еще один вариант: сосредоточиться не на частоте или интенсивности, а на временных характеристиках поведения. Исследователь, к примеру, может решить измерять *латентность* или скорость агрессивной реакции или же общую *продолжительность* эпизодов нанесения ударов. Эта трихотомия «частота—интенсивность—продолжительность» применима не ко всем видам поведения, изучаемым в психологии развития; однако во многих случаях ее вполне можно использовать. Если же такая возможность отсутствует, обычно имеются другие параметры, которые поддаются измерению. Редко какой результат подразумевает лишь единственный способ его измерения.

Выбор аспекта поведения для измерения предполагает наличие решения о способе измерения и, следовательно, ограничен внешними поведенческим проявлениями. Объектом измерения не всегда становится внешнее поведение. Исследователи агрессии могут интересоваться и агрессивно окрашенные мысли или фантазии, то есть подспудное психическое содержание, а не фактическое поведение. Чтобы сделать вывод о психическом содержании, ему все равно необходимо выделить некую измеряемую реакцию (например, самоотчеты об агрессивно окрашенных фантазиях); однако в этом случае поведение — это лишь средство к достижению цели. Аналогично исследователь агрессивных эмоций должен получить вербальные отчеты об этих эмоциях; однако вновь объектом изучения служит нечто отличное от внешних поведенческих проявлений. В качестве альтернативы исследователь мог бы выбрать обходной путь и измерять не поведенческие реакции, а физиологические показатели (например, повышение частоты сердечных сокращений, повышение артериального давления) в ответ на ситуацию, провоцирующую агрессию. Эмоции — это только один из конструктов, которые позволяют производить не выраженные внешне измерения. В дальнейшем мы еще встретим множество примеров подобных конструктов.

Обратите внимание, что описанные в предыдущем абзаце варианты оценки иллюстрируют различные направления измерений. Одно из них — от явного к внешне не выраженному. В одних случаях нас интересует само поведение, в других — не наблюдаемые, более общие конструкты (мысли, мотивы, потребности и т. д.), которые предположительно лежат в основе видимого поведения. В одних случаях специфический объект наших измерительных операций — явное поведение (например, удары), в других — это некий скрытый недоступный для прямого наблюдения процесс (например, частота сердечных сокращений). Наконец, в некоторых случаях имеющееся в наличии и специфическое совпадает, в других — нет. Если говорить об ударах, то здесь присутствует совпадение — нас интересует, как часто ребенок бьет других детей, и мы измеряем частоту ударов. Исследователя же эмоций, скорее всего, не интересует частота сердечных сокращений сама по себе; частота сердечных сокращений — это лишь ключ к эмоциональному состоянию. Это различие определяют как различие между *признаком* и *специфической моделью*. Иногда мы измеряем специфические модели конструкта, то есть частные случаи (удары, плач, улыбки и т. д.) интересующего нас поведения. А иногда — только признаки, не конструкт сам по себе (поднятие бровей, изменение частоты сердеч-

ных сокращений и т. д.), по которым можно судить об этом конструкте. И разумеется, одна и та же характеристика может выступать либо в роли признака, либо в роли специфической модели, в зависимости от ее функции в конкретном исследовании. При изучении плача приступы плача, естественно, — специфические модели. При изучении привязанности приступ плача это признак — обычно один из множества, — по которому судят о характере привязанности.

Последнее разграничение касается целей измерения, то есть того, что мы собираемся делать с полученными данными. В некоторых случаях цель — выявление индивидуальных различий между испытуемыми, оценка степени различия испытуемых из одной выборки по параметру агрессии, привязанности или чего-либо еще. Такая цель особенно часто стоит перед корреляционным исследованием, в котором осуществляется попытка выявления связи между вариациями значений одной группы показателей (например, индивидуальных различий в уровне агрессии у детей) и вариациями значений другой группы показателей (например, различий в полученном воспитании). В других исследованиях в центре внимания находятся не столько индивидуальные различия, сколько непосредственные детерминанты изучаемого поведения. Возьмем в качестве примера гипотетическое исследование социального подкрепления и агрессии. Цель такого исследования могла бы состоять в определении отношений между колебаниями уровня агрессии и наличием или отсутствием подкрепления, а не в оценке того, кто из детей более или менее агрессивен. Или же рассмотрим экспериментальное изучение связи просмотра телепрограмм с эпизодами насилия и агрессивного поведения, подобное тому, которое провели Либерт и Бэрон (Liebert & Baron, 1972). Вновь цель не в выявлении индивидуальных различий в уровне агрессии; цель — узнать, изменяется ли уровень агрессии как функция от характера телепрограмм, которые смотрят дети.

Описанное выше различие иногда называют различием между *свойством* и *состоянием*. При измерении свойства нас интересует вопрос, каковы люди вообще, при этом цель, как правило, состоит в выявлении связи измеряемой характеристики с неким другим показателем или некими другими показателями в одной и той же выборке. При измерении состояния нас интересует вопрос, каковы люди *в данный момент*, при этом цель, как правило, состоит в выявлении связи вариаций непосредственных поведенческих реакций с некоей потенциальной детерминантой изучаемого поведения. В обоих случаях, несомненно, важно выбирать адекватную операционализацию изучаемого конструкта. Более того, конкретные измерительные операции могут быть общими и для оценки состояния, и для оценки свойства. Если, к примеру, определенный итоговый индекс физических актов (ударов, пинков, щипков и т. д.) выступает как критерий индивидуальных различий по уровню агрессии, то этот же критерий, вероятно, может быть использован и в экспериментальном исследовании эффектов просмотра программ с эпизодами насилия. Однако, даже если критерии совпадают, способы измерения этих параметров и их использования могут быть разными в разных видах исследования. Когда в центре внимания находятся индивидуальные различия между людьми, первостепенное значение приобретает оценка специфических моделей поведения в выборке: возможность упорядочить результаты испытуемых по определенному параметру за-

висит от возможности оценить репрезентативные модели поведения каждого испытуемого. Когда же в центре внимания находится влияние экспериментальных манипуляций на изучаемое поведение, вопрос измерений в выборке, а также индивидуальные различия отходят на второй план: теперь нас скорее всего будет интересовать общий, сравнимый с другими, тип поведения, который позволяет выявить искомый эффект. В действительности в этом случае имеющиеся индивидуальные различия становятся даже помехой, поскольку вносят дисперсию ошибки, которая может исказить эффект, представляющий исследовательский интерес.

В этом разделе анализировался ряд направлений, по которым различаются измерения. Все они сведены в табл. 4.1. Как уже отмечалось, более полное освещение этой темы (например, Messick, 1983) добавило бы ряд новых граней. Однако несмотря на это многообразие параметров измерения, принцип, который следует уяснить, довольно прост: мы выбираем стратегии измерения согласно конкретным целям исследования.

Таблица 4.1

Аспекты измерения

Направление	Описание
Аспект поведения	Какой из аспектов поведения (например, частота, интенсивность, временные параметры) измеряется
Явное/внешне не выраженное	Явный объект измерения (наблюдаемое поведение) или неявный (например, некие физиологические изменения)
Признак/специфическая модель	Является ли критерий признаком, по наличию которого судят о конструкте, или специфической моделью данного конструкта
Свойство/состояние	Находятся в центре внимания устойчивые индивидуальные различия между испытуемыми или непосредственный эффект экспериментальных манипуляций

Качество измерения

Как было сказано выше, измерение — это процесс принятия решений: выбора ограниченного количества методов оценки конструкта из значительно большей совокупности методов. В этом разделе мы рассмотрим некоторые факторы, определяющие, насколько удачен выбор того или иного способа измерения.

Как говорилось в главе 2, «изменчивость» зависимой переменной подразумевает, что получаемые результаты с некоторой долей вероятности изменяются как функция от факторов, изучаемых исследователем. Поэтому один из критериев качества измерения состоит в том, чтобы оно давало необходимый уровень и диапазон значений переменной. В противном случае показатели будут так скучены, что ни один эффект не проявится. Чаще всего этот вопрос обсуждают, говоря о корреляционном исследовании и проблеме ограничения диапазона переменных (смот-

ри главы 3 и 7); однако он имеет отношение и к эксперименту. Рассмотрим вновь исследование Черри и Парка (Cherry & Park, 1993), описанное в главе 2. Предположим, что в качестве стимульного материала для оценки памяти они предъявляли бы не 24, а 4 предмета. В этом случае большинство испытуемых, вероятно, безупречно справились бы с заданием, и не было бы никакой возможности обнаружить эффект контекста (схемы или модели), на выявление которого было направлено исследование.

Эта гипотетическая переформулировка исследования Черри и Парка иллюстрирует один из возможных видов ограничения диапазона: так называемый эффект верхнего предела. Об **эффекте верхнего предела** говорят, когда задание слишком легкое, и поэтому большинство испытуемых получает баллы, находящиеся на верхней границе шкалы или очень близко от нее. Возможна и противоположная ситуация — скученность показателей у нижней границы шкалы вследствие слишком высокого уровня трудности заданий. Эта проблема, как вы, наверное, догадались, называется **эффектом нижнего предела**.

При проявлении достаточной предусмотрительности обычно удается избежать эффектов и верхнего, и нижнего пределов. Поиску оптимального уровня трудности может помочь предварительное исследование с аналогичными заданиями и выборкой. Если сомнения все же остаются, для уточнения критериев можно провести пилотажное исследование. Но как бы ни был предусмотрителен исследователь, он все равно может столкнуться с трудностями, если одни и те же задания и процедуры должны использоваться с испытуемыми, находящимися на разных уровнях развития. Задание, умеренно трудное для одной возрастной группы детей, может вызвать эффект верхнего предела детей старшего возраста, эффект нижнего предела — у детей младшего возраста.

Обсуждая подбор заданий для разных возрастных групп, мы затрагиваем глобальный вопрос *эквивалентности измерения*. Как говорилось в главе 3, проблема эквивалентности измерений становится актуальной всякий раз, когда в исследовании сравниваются группы испытуемых. В психологии развития наиболее часто этот вопрос встает при сравнении возрастных групп, однако может иметь значение и при сравнении любых групп, сформированных естественным путем. При изучении различий между полами по уровню когнитивных способностей, к примеру, важно отбирать материал, одинаково знакомый и интересный как для мальчиков, так и для девочек. При сравнении разнообразных этнических и культурных общностей важно — и зачастую весьма затруднительно — подобрать методики, которые подходили бы для представителей всех этих групп. Довольно трудно бывает подобрать критерии, адекватные для одной группы испытуемых. Трудность измерения возрастает в несколько раз, когда цель исследования состоит в выявлении межгрупповых различий.

Из сказанного об эквивалентности измерений следует, что одного типа операционализации изучаемого конструкта зачастую бывает недостаточно. Методика, пригодная для одной группы, может оказаться непригодной для другой группы. Это явление — частный случай общей закономерности: Даже когда не производится сравнение групп, результаты использования лишь одной методики — неустойчивый базис для получения выводов. Кук и Кэмпбелл (Cook & Campbell, 1979) на-

зывают применение только одного типа измерительных операций **монооперационным искажением**. По их мнению, одна операция практически всегда дает недостаточное представление об изучаемом конструкте, то есть раскрывает лишь некоторые аспекты того (агрессий, интеллекта, половой типизации и т. д.), что мы пытаемся измерить. Наряду с этим любое измерение включает ряд специфических для каждого задания дополнительных факторов (например, определенный порядок слов в инструкции, характер стимульного материала, конкретный человек, выполняющий функции экспериментатора), и поэтому трудно с помощью только одного типа измерений узнать, какая часть полученных результатов действительно имеет значение, а какая — следствие действия дополнительных факторов. Можно быть гораздо более спокойными за наши выводы при использовании нескольких форм измерения. Этот принцип вам уже знаком; в главе 3 мы говорили то же самое об опасности монометодического подхода при изучении причинно-следственных связей. Правилom для обоих случаев является использование *конвергентных операций*, а не единственного метода исследования.

В этом разделе, посвященном качеству исследования, остается рассмотреть еще два конструкта, крайне значимых для оценки измерения: надежность и валидность. **Надежность** характеризует согласованность или воспроизводимость измерений. Вопрос состоит в следующем: будут ли результаты повторных применений некоей методики одинаковы или весьма близки? Чем выше степень согласованности при неоднократном применении, тем выше надежность. Цель измерения всегда заключается в максимизации надежности. *Валидность*, как и везде, имеет отношение к точности — в данном случае, к точности измерений. Вопрос состоит в том, точно ли отражают полученные в ходе измерения значения изучаемый конструкт, то есть измеряем ли мы именно то, что предполагали измерять? Валидность, очевидно, вопрос, касающийся всего процесса измерения. Это глобальная цель, охватывающая все частные моменты, обсуждавшиеся в данном разделе.

Всегда, когда мы что-либо измеряем, встают вопросы надежности и валидности. Однако на практике эти конструкты чаще всего рассматриваются в их отношении к двум видам измерения: к стандартизованным тестам и к обсервативным оценкам. Стандартизованным тестам и методу наблюдения посвящены оставшиеся разделы этой главы; в контексте этих двух форм измерения я более подробно остановлюсь на вопросах надежности и валидности.

Тесты

Для определения, достойна ли некая измерительная операция называться тестом, не существует общепринятых критериев. В том смысле, в котором данный термин употребляется здесь, это измерительный инструмент, выполняющий функцию оценки ряда важнейших психологических характеристик. Существует множество психологических характеристик и еще больше тестов, претендующих на их измерение; список их включает несколько тысяч наименований (например, Bugos, 1978). На страницах этой книги вы встретите ряд примеров: методики оценки темперамента в младенчестве (глава 10), Я-концепции у дошкольников (глава 12), полоролевого развития у старших детей (глава 12) и интеллекта в разные периоды

жизни (главы 10, 11 и 13). Приведенные в качестве примера тесты — как и все тесты вообще — различаются по объекту оценки, по возрастным группам, для которых они предназначены, по схеме получения ответов и способам использования результатов тестирования. Однако общим для них всех является то, что они должны отвечать двойному критерию — валидности и надежности.

Валидность

При оценке **валидности теста** ставится следующий вопрос: измеряет ли тест то, что, как предполагается, он должен измерять? Если это, к примеру, *IQ*-тест, действительно он измеряет различия в уровне интеллекта или различия показателей испытуемых обусловлены чем-то другим? Очевидно, что недостаточно назвать тест тестом интеллекта, нужны другие критерии. В целом, существует три типа критериев валидности.

Первый — тест должен отвечать требованиям **содержательной валидности**. Термин *содержательная валидность* означает адекватность, с которой задания теста характеризуют изучаемое понятие. Включает ли тест все важнейшие аспекты объекта, которые мы хотим измерить, и правильно ли оценивается значение каждого из них? Предположим, наш тест должен оценивать знания арифметики на уровне 4-го класса. Тест, содержащий примеры только на сложение, будет обладать недостаточной содержательной валидностью. Значительно лучшей содержательной валидностью будет обладать тест, включающий репрезентативную подборку задач на сложение, вычитание, умножение и деление.

Добиться содержательной валидности желательно, но это не всегда легко. Даже в отношении такого строго структурированного предмета, как элементарная арифметика, могут возникнуть расхождения, касающиеся адекватности подобранных заданий. К примеру, сколько должно быть примеров с двузначными, а сколько — с трехзначными числами; в какой контекст или в какие контексты должны быть помещены эти примеры? Когда объектом является нечто более сложное, чем элементарная арифметика, добиться содержательной валидности бывает практически невозможно. Как бы ни был широк диапазон подобранных вопросов теста, сомнительно, чтобы с помощью одного контент-анализа можно было бы продемонстрировать полноту и репрезентативность подбора вопросов для оценки любого из возможных аспектов «интеллекта». В таких случаях необходимы другие индексы валидности.

Вторая форма валидности — **критериальная валидность**. Здесь вопрос состоит в том, связаны ли результаты выполнения теста испытуемым с определенным внешним критерием характеристики, оцениваемой этим тестом. В случае теста на знание арифметики разумным критерием были бы отметки по арифметике за учебный год. Тест, значительно коррелирующий с этими отметками, обладал бы необходимой критериальной валидностью. В случае с интеллектом характерным критерием является возможность с помощью *IQ*-тестов спрогнозировать уровень школьной успеваемости или успешность выполнения стандартизированных тестов достижения; в действительности, исторически именно потребность в инструментах прогноза школьной успеваемости заставила приступить к разработке первого *IQ*-теста (тест Симона—Бине, 1905 год). Чаще всего критериальная валидность — это главная форма валидности тестов, основная функция которых состоит

в получении прогностической прагматической оценки. Поэтому именно критериальная валидность лежит в основе использования тестов академических способностей (Scholastic Aptitude Test — SAT) или письменных экзаменов для выпускников (Graduate Record Examinations — GRE) с целью прогноза успеваемости в колледже; использования результатов забега на 40 ярдов с целью прогноза успехов профессиональных футболистов и т. д.

Выделяют две формы критериальной валидности. Тест, коррелирующий с неким наличествующим в данный момент внешним критерием, считается обладающим *текущей валидностью*. Доказав, к примеру, что *IQ*, измеряемый во втором классе, коррелирует со школьной успеваемостью во втором классе, мы продемонстрировали бы текущую валидность нашей методики оценки *IQ*. О тесте, коррелирующем с неким внешним критерием *в будущем*, говорят, что он обладает *прогностической валидностью*. Доказав, что *IQ*, измеренный во втором классе, связан с успеваемостью в старших классах, мы продемонстрировали бы прогностическую валидность.

Последняя форма валидности, которую мы рассмотрим, называется **конструктивной валидностью**. Среди психометристов конструктивная валидность считается наиболее важной формой валидности теста. К сожалению, из всех видов валидности исследователю ее труднее всего добиться, а автору учебника — труднее всего разъяснить. Я ограничусь лишь краткими предварительными замечаниями об этом сложном понятии. Более подробное изложение данной темы можно найти во многих источниках, включая работы Кронбах (Cronbach, 1990), Керлингера (Kerlinger, 1986) и Наннэлли (Nunnally, 1978).

Характерной особенностью конструктивной валидности является ее теоретическая направленность. Как замечает Керлигер (Kerlinger, 1986), «это не просто вопрос валидации теста. Необходимо попытаться валидизировать теорию, стоящую за тестом» (р. 420). Исходным пунктом поэтому становится некая концепция конструкта (интеллекта, креативности, Я-концепции, тревожности и т. д.), который мы хотим измерить. На основе этой концепции можно выдвинуть ряд гипотез. Это могут быть гипотезы об эффекте определенного рода экспериментальных манипуляций. Предположим, что мы пытаемся валидизировать некую методику измерения тревожности. Можно выдвинуть гипотезу о том, что повышение напряженности тестовой ситуации приведет к повышению показателя, полученного при помощи данной методики, и наоборот, снижение напряженности приведет к снижению этого показателя. Подобные результаты явились бы свидетельством наличия конструктивной валидности.

Для установления конструктивной валидности, помимо экспериментальных тестов, имеют значение и корреляционные данные. Прогнозируемые корреляции бывают двух видов. Одни гипотезы касаются вопроса о том, какие параметры конструкта положительно коррелируют с некими другими параметрами. Разрабатывая, к примеру, тест тревожности, мы могли бы предположить, что самоотчеты о переживании чувства тревоги коррелируют с физиологическими изменениями, которые рассматриваются как индикаторы тревоги (например, повышение частоты сердечных сокращений). Наличие прогнозируемой корреляции между показателями, имеющими общую теоретическую основу, называется *конвергентной ва-*

Метод	Свойство	
	A	B
	A1	B2
	A2	B2
1	A1	B2
2	A2	B2
3	A3	B3

Рис. 4.1. Пример матрицы «свойства × методы».

Буквами обозначены свойства или конструкты, а числами — методы исследования

лидность. Другие гипотезы касаются вопроса о том, какие показатели *не* должны коррелировать друг с другом. Например, при валидизации теста тревожности необходимо доказать, что некоторые физиологические изменения не связаны с самоотчетами о переживании чувства тревоги, таким образом исключая общее возбуждение как объяснение полученных результатов. Дифференциация показателей, имеющих разную теоретическую основу, называется *дивергентной (или дискриминантной) валидностью*.

Весьма оригинальное решение вопроса конструктивной валидности предложили Кемпбелл и Фиске (Campbell & Fiske, 1959), разработавшие **матрицу «свойства × методы»** (СМ). Использование матрицы СМ подразумевает анализ корреляций между предполагаемыми свойствами или конструктами, каждый из которых измеряется при помощи разных методов. В примере, проиллюстрированном на рис. 4.1, исследуется три свойства, каждое из которых оценивается при помощи трех методов, результатом чего является 9 показателей (A1, A2 и т. д.), составляющих массив таблицы. Конвергентную валидность демонстрировали бы достаточно высокие коэффициенты корреляции между показателями каждого столбца, то есть показателями, которые, как предполагается, характеризуют один и тот же конструкт и отличаются только методами, посредством которых они были получены. Дивергентную валидность демонстрировали бы низкие коэффициенты корреляции между любыми показателями, не попадающими в один столбец. Особенно интересны в этом отношении корреляции между показателями каждой строки — то есть, показателями, полученными при использовании одного и того же метода оценки. Иногда положительные корреляции между результатами обусловлены только методическим сходством заданий, а не характером измеряемых параметров. К примеру, в ряде тестов требуется быстрая реакция в условиях ограничения времени; тогда, что бы ни измеряли тесты, испытуемые, чувствуя себя в такой обстановке более уверенно, справятся с заданиями лучше. О корреляциях, являющихся результатом частичного совпадения методов, говорят, что они отражают **дисперсию общности методов**. Матрица СМ позволяет определить вклад дисперсии общности методов во все полученные корреляционные связи.

Как следует из вышесказанного, конструктивная, как и критериальная валидность в большинстве случаев оценивается через проверку ожидаемых корреляций между результатами измерений. Однако между этими двумя видами валидности имеются существенные различия. Критериальная валидность обычно оценивается по какому-то одному внешнему показателю, например школьной успеваемости, в отношении которого мы хотим сделать прогноз; конструктивная же валидность оценивается по целой системе предполагаемых взаимосвязей. Цель определения кри-

териальной валидности, как правило, — в прагматическом прогнозе; цель определения конструктивной валидности — валидизация лежащей в основе теста теории. Поэтому то, что последняя из рассмотренных форм валидности носит то же название, что и одна из форм валидности эксперимента, о которой рассказывалось в главе 2, не случайно (хотя и может смутить читателя). В обоих случаях сутью вопроса является теоретическая обоснованность: в одном — в отношении измерения, в другом — в отношении исследования в целом.

Надежность

Стандартизованный тест должен обладать не только валидностью, но и удовлетворять критерию надежности. Вопрос надежности в его применении к тестам также довольно очевиден: Согласованы ли результаты измерения данным тестом? Предположим, что мы даем одному ребенку *IQ*-тест несколько раз подряд, а затем сравниваем результаты. Если они близки, тест обладает хорошей надежностью; значительный разброс результатов свидетельствовал бы о недостаточной надежности.

Пример с *IQ* иллюстрирует одну из основных форм надежности — **ретестовую надежность**. Существует два способа оценки ретестовой надежности. Один — дать один и тот же тест дважды. Однако понятно, что если тесты одинаковы, ребенок может вспомнить свои ответы, а это приведет к искусственному завышению надежности (это также может привести к *занижению* надежности, если ребенок воспримет повторное предъявление теста как сигнал к изменению своих ответов). Дабы избежать этой проблемы, ретестовую надежность иногда оценивают с использованием *взаимозаменяемых форм теста*. Как следует из названия, такой подход требует наличия двух разных, но эквивалентных версий теста, при этом один вариант предъявляется в момент 1, а другой — в момент 2. Вновь высокая согласованность ответов будет свидетельствовать о высокой надежности.

Второй из основных типов надежности называется **надежностью внутренней согласованности**. Теперь суть состоит в согласованности ответов на разные вопросы одного теста, предъявляемого однократно. Как правило, для этого пользуются процедурой «расщепления» теста на нечетные и четные задания с последующим сравнением ответов между двумя этими категориями. Вновь о высокой надежности будет говорить высокая согласованность ответов.

Важно отличать надежность от других конструктов, которые также имеют отношение к согласованности независимо от полученных результатов. Предположим, что интервал между первым и вторым выполнением теста интеллекта не один день, а два года. Обнаружив значительное расхождение между этими двумя показателями, должны мы сделать вывод о ненадежности теста или о том, что *IQ* ребенка действительно изменился за два года? Допустим, мы решили измерять не интеллект, а вес. Если наши измерения показывают, что 9-летний ребенок весит на 15 фунтов больше, чем он весил в 7-летнем возрасте, должны ли мы заключить, что наши весы неисправны? Разумеется, вероятнее всего, что наш ребенок действительно прибавил в весе за прошедшие два года, то есть в период физического развития стабильность веса далека от идеальной. Многие аспекты жизнедеятельности ребенка (включая успешность выполнения *IQ*-тестов) далеки от идеальной стабильности

в период его развития. Поэтому важно понимать разницу между надежностью измерения и стабильностью поведения.

Надежность измерения также важно отличать от типичности поведения. Вопрос типичности — это вопрос согласованности поведения в разных ситуациях. Предположим, нас интересует агрессия у дошкольников. Мы идем в детский сад и регистрируем все наблюдаемые там акты агрессии. На основе полученных данных мы вычисляем индекс агрессии для каждого ребенка. Затем мы отправляемся домой к детям и измеряем агрессию там. Обнаруживается, что наши показатели агрессии в домашних условиях слабо связаны с показателями агрессии в детском саду — иными словами, эти показатели существенно расходятся. Заключим ли мы, что один или оба метода измерения ненадежны? Хотя это вполне возможный вывод, более разумно было бы заключить, что уровень агрессии зависит от обстановки. В этом случае обнаруженный факт связан с типичностью поведения, а не с надежностью измерений.

Вероятно, полезно было бы резюмировать сказанное выше. Надежность — это свойство метода измерения; стабильность и типичность — это свойства поведения. Надежность — это то, что исследователь всегда пытается максимизировать. Стабильность и типичность же — это феномены, которые нужно изучать, а не максимизировать. Наконец, эти феномены *можно* изучать только по достижении достаточно высокого уровня надежности. Только будучи уверенными в надежности наших методов для определенного времени и ситуации, мы можем попробовать оценить согласованность поведения во времени (вопрос стабильности) и в разных ситуациях (вопрос типичности).

Регрессия

Вернемся на время к ретестовой надежности. Надежность, которая далека от идеальной подразумевает, что результаты второго выполнения теста отличаются от результатов, полученных при первом тестировании. Можно ли уточнить это обобщенное утверждение и сказать что-либо о *направлении* изменения, то есть повысятся или понизятся результаты при втором предъявлении? Если рассматривать индивидуальные случаи, то предугадать направление измерения нельзя. Однако на уровне среднegrupповых показателей такой прогноз *возможен*. В целом испытуемые, получающие при первом тестировании низкие результаты, во второй раз справляются с заданиями успешнее, а испытуемые, получающие при первом тестировании высокие результаты, на второй раз справляются хуже. Тенденция, проявляющаяся в изменении при повторном тестировании изначально крайних показателей в направлении группового среднего, называется **регрессией к среднему**.

Прежде чем попытаться объяснить, почему это происходит, рассмотрим конкретный пример. Допустим, мы предлагаем отобранной группе детей некий тест *IQ* и получаем распределение баллов, изображенное на рис. 4.2. Часть детей (пустые кружочки) получила оценки значительно ниже среднего, часть (закрашенные кружочки) — значительно выше среднего, а результаты остальных (зачеркнутые кружочки) находились в пределах среднего. Теперь предположим, что мы даем тот же тест той же группе неделю спустя и получаем распределение, изображенное на рис. 4.3. Как можно заметить, у детей, получивших первоначально низкие резуль-

таты, успешность в среднем повысилась, а у детей, первоначально получивших высокие результаты — снизилась. Таким образом, у обеих групп наблюдалась регрессия к среднему. Однако, поскольку некоторые из тех, кто первоначально справился с заданиями на среднем уровне, впоследствии улучшили или ухудшили свои результаты, общий диапазон значений и среднее значение остались неизменными.

Почему происходит регрессия? Результат любого испытуемого можно рассматривать как сумму двух компонентов — «истинного результата», или фактического значения измеряемого параметра, и погрешности, обусловленной несовершенством измерения. Ясно, что «погрешность» — это еще один способ описания надежности: абсолютная надежность подразумевает отсутствие погрешностей; и наоборот, чем больше погрешность, тем ниже надежность. В отношении погрешности обычно бывают оправданны два допущения. Первое — среди истинных показателей погрешность распределена по закону нормального распределения. Это означает, что небольшие погрешности встречаются чаще, чем существенные; это также означает, что погрешности с равной вероятностью могут привести как к завышению, так и к занижению индивидуальной оценки каждого испытуемого. Второе допущение — погрешности случайным образом распределены между испытуемыми и между тестовыми ситуациями. Это означает, что погрешность в результатах определенного испытуемого в одном тесте не коррелирует с погрешностью этого же испытуемого в другом тесте; погрешность в первом тесте не связана погрешностью во втором тесте.

Рассмотрим теперь результаты, изображенные на рис. 4.2. Каким образом на них отразились погрешности измерения? В частности, как повлияли на полученное распределение существенные погрешности, то есть те, которые приводят к значительному завышению или занижению результата? Справедливо допустить, что низкие показатели (пустые кружочки), в целом, явились следствием непропорционального количества отрицательных погрешностей; это одна из причин низких результатов части детей. Аналогично справедливо допустить, что высокие показатели, в целом, явились следствием непропорционального количества положительных погрешностей; это одна из причин высоких результатов другой части детей. Но что происходит при повторном тестировании? Вспомним, что погрешности результатов одного тестирования не связаны с погрешностями результатов другого. Поэтому маловероятно, что существенные погрешности изменят результаты тех же детей и в том же направлении. Скорее всего, результаты каждого ребенка затронет относительно небольшая погрешность, которая равно вероятно может исказить истинный показатель, как в сторону повышения, так и в сторону понижения. Это «выравнивание» погрешностей при повторном тестировании обуславливает тенденцию низких показателей к повышению, а высоких — к понижению; иными словами, обуславливает регрессию к среднему.

Основная проблема, которую ставит регрессия перед исследователем, очевидна. Как и недостаточная надежность в целом, регрессия представляет угрозу для валидности исследования. Более того, поскольку регрессия — систематический феномен, она может явиться причиной систематически неверных выводов. Предположим, что в описанном ранее исследовании *IQ* мы не только повторно протестировали детей, но и включили между первым и вторым тестированием новую

образовательную программу. Получив результаты, указанные на рисунках 4.2 и 4.3, мы могли бы заключить, что эффект нашей программы зависит от изначального уровня способностей: программа привела к повышению *IQ* слабоодаренных детей, но фактически снизила *IQ* у одаренных. Очевидно, что в подобном случае регрессия может произвести ложное впечатление изменения. Или же регрессия может маскировать истинное изменение; к примеру программа действительно развивает у детей способности, однако истинное улучшение показателей делается незаметным из-за потерь, обусловленных регрессией.

Исследования программ вмешательства, таких как описанные выше, являются наиболее типичным контекстом проявления эффекта регрессии, так как выборкой для них обычно служат дети с низкой успеваемостью. Этому недостатку подвержены также определенного вида планы уравниваемых групп. Рассмотрим несколько измененную схему описанного в главе 3 исследования лиц, окончивших и не окончивших школу. Представьте теперь, что вас интересует стабильность *IQ* во времени. Сохраняются ли способности у окончивших школу лучше, чем у не окончивших ее? Вы уравниваете группы по показателю *IQ*, среднему между показателями обеих групп (скажем, 97) и проводите повторное тестирование 10 лет спустя. Принимая во внимание лишь регрессию, мы можем предположить, что средний показатель тех, кто окончил школу, повысится (поскольку отобраны представители популяции, имевшие относительно низкие результаты), а средний показатель тех, кто не окончил школу, снизится (поскольку отобраны представители популяции, имевшие относительно высокие результаты). Вновь регрессия может произвести заметный эффект, который не имеет ничего общего с эффектом изучаемой независимой переменной.

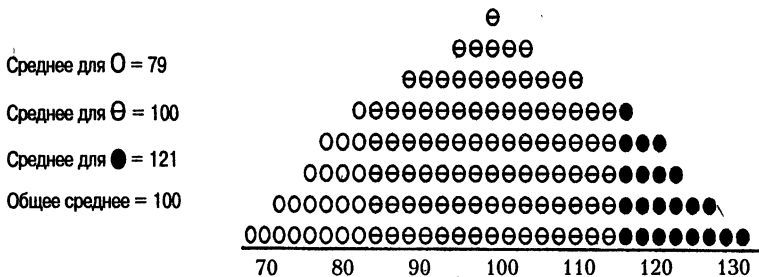


Рис. 4.2. Гипотетическое распределение результатов при первом предъявлении теста *IQ*

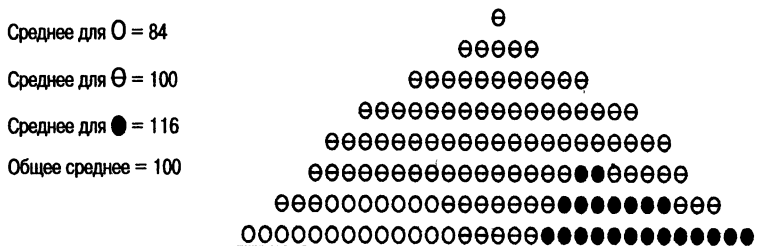


Рис. 4.3. Гипотетическое распределение результатов при повторном предъявлении теста *IQ*

Методы наблюдения

Прямое наблюдение поведения одновременно и наиболее ценный и наиболее сложный из методов психологического исследования. Поэтому я завершу эту главу обобщением некоторых трудностей, связанных с наблюдением поведения.

Сначала нужно внести некоторую ясность. В определенном смысле, всякое исследование включает наблюдение поведения: как еще можно было бы оценить значение зависимой переменной? В ряде случаев регистрация поведения происходит практически, если не буквально, автоматическим образом. Реакции со стороны сердца могут фиксироваться на электрокардиограмме. Решая задачи, испытуемый может давать ответы, нажимая на кнопку. Достаточно взрослым испытуемым предлагаются опросники. Каковы бы ни были другие проблемы в таких исследованиях, проблема точности регистрации поведенческих реакций, как правило, не стоит.

При наблюдении точность регистрации определенно является проблемой. Обсервативное исследование обычно направлено на изучение довольно продолжительных эпизодов естественного поведения, которое невозможно зарегистрировать автоматически; от наблюдателя требуется вынесение суждений относительно наличия или отсутствия определенного поведения и о его значении. Поэтому центральный вопрос состоит в том, какова объективность решения наблюдателя. Разделим тему методов наблюдения на три основные проблемы: что наблюдать, как наблюдать и как определить точность результатов наблюдения. Более подробно с этими вопросами можно ознакомиться у Хартман и Вуд (Hartman & Wood, 1990), Одом и Огава (Odom & Ogawa, 1992), Сэккетт (Sackett, 1978), а также Йарроу и Уакслер (Yarrow & Waxler, 1979).

Что наблюдать

На определенном уровне ответы на вопрос «что» довольно очевидны. Понятно, что общие интересы исследователя ограничены поведением, которое можно наблюдать. Характер поведения, в свою очередь, определяет, насколько целесообразна та или иная стратегия наблюдения. Одни формы поведения легче изучать методом наблюдения, чем другие. Агрессия, к примеру, естественный претендент на обсервативную оценку: частое, наблюдаемое, «зримое» поведение. Хотя существуют и другие способы измерения (например, рейтинговые шкалы, изолированные экспериментальные тесты), они в меньшей степени отвечают целям исследования. И наоборот, частота сердечных сокращений и физиологические реакции не подходят в качестве объекта обсервативной оценки. Такие реакции трудно, а зачастую невозможно увидеть непосредственно, и применение других методов будет и более легким, и более разумным выбором.

Однако ситуация осложняется, когда исследователь выходит за рамки первоначального решения использовать методы наблюдения, пытаясь определить, какие именно аспекты поведения следует регистрировать. Предположим, мы изучаем стиль взаимодействия матери со своим младенцем. Для начала мы должны осознать, что все запротоколировать невозможно; наблюдение всегда связано с некоторым абстрагированием от конкретики каждого момента. Но до какого уровня конкретности следует абстрагироваться? Нужно ли отмечать тот факт, что мать

подняла брови, широко раскрыла глаза, приподняла уголки рта, издала звук. Или мы должны работать на более глобальном, интерпретационном уровне, отмечая, что мать улыбнулась и заговорила с малышом? Или использовать еще более обобщенную систему интерпретации и указать на то, что мать поощряет только что произведенные действия младенца? Или мы должны перейти на еще более высокий уровень и отметить, что мать позитивно и доброжелательно относится к своему ребенку?

Проведенные выше разграничения обозначают как оппозицию микро- и макронаблюдения (Sackett, Ruppenthal & Gluck, 1978). Использование **микросистемы наблюдения** предполагает пристальное внимание к тонким деталям поведения, максимальную приближенность к фактическому поведению и практически нейтральное, четкое описание происходящего. Разумеется, некоторые детали все-таки опускаются и присутствует некоторая интерпретация; но даже в этих условиях целью является по возможности наиболее полное, конкретное и безоценочное описание. **Макросистема наблюдения**, напротив, предполагает некоторый отход от фактического описания, суммирование микроэлементов наблюдения, результатом которого становится формулирование определенной оценочной категории. «Улыбка» или «объятия» — примеры макрокатегорий на относительно специфическом уровне; «поощрение» и «отвлечение» — примеры более глобального, интерпретационного уровня.

Как следует из вышесказанного, противопоставление микро- и макросистем предполагает не столько дихотомию, сколько наличие континуума. В разных системах наблюдения может реализовываться разное соотношение конкретики и интерпретации. Таблицы 4.2 и 4.3 являют собой пример систем взаимодействия между матерью и младенцем, систем, которые довольно точно попадают в категории микро- и макронаблюдений. В системе, разработанной Элс, Троник и Бразелтон (Als, Tronick, & Brazelton, 1979) (табл. 4.2), анализируются наиболее элементарные формы адаптации младенца при взаимодействии с матерью. В системе, разработанной Лэмб (Lamb, 1976) (табл. 4.3), анализируются способы поддержания младенцем контакта со своими родителями в незнакомой ситуации, включая возможные различия в реагировании на мать и на отца. Главными направлениями оценки являются две глобальные макрокатегории: дистальное/аффилиативное поведение, проксимальное поведение/поведение привязанности.

То, где скорее всего будет работать исследователь на континууме от микро- к макросистемам, зависит от двух основных факторов. Один из них — цель исследования. Если она заключается в выявлении мимических признаков разнообразных эмоциональных состояний (например, Izard, 1979), естественно, необходима микросистема наблюдения. Если же цель в изучении терминант улыбки или смеха (например, Sroufe, Waters, & Matas, 1974), более разумно использовать макросистему. Конечно, в одном исследовании можно сочетать микро- и макроанализ, особенно если поведение фиксируется на видеопленке. Однако обратите внимание на однонаправленность перехода от одной системы к другой при использовании письменной регистрации. Как правило, имеется возможность, применив микросистему, затем перевести полученные результаты в макроинформацию — например, провести анализ мимики с последующим определением частоты улыбок. Используя же макросистему, невозможно перевести результаты в микроинформацию.

Вторую из главных детерминант уровня наблюдения можно обозначить как осуществимость, что бы исследователь ни хотел наблюдать, решающим условием будет *возможность* наблюдения. Регистрация микрохарактеристик, к примеру, может оказаться осуществимой только при тесном контакте с испытуемыми или только при наличии видеокамеры. Исследователи, работающие в других условиях (к которым, разумеется, относятся максимально естественные), бывают вынуждены ограничиться макросистемой наблюдения. Еще одним фактором, о котором вскоре мы поговорим более подробно, является надежность методов наблюдения. Результаты наблюдения не будут иметь смысла, если два независимых наблюдателя расходятся во мнении относительно наблюдаемого явления. Иногда требование микросистемы регистрировать мельчайшие подробности превосходит возможности наблюдателей, вынуждая исследователя перейти к более обобщенным категориям оценки. Иногда необходимость интерпретации при использовании макросистемы (были действия матери проявлением отвержения или просто попыткой переключить внимание ребенка?) может приводить к постоянным разночтениям, вынуждая исследователя перейти на уровень, более близкий к фактическому (например, отвернулась от ребенка). В чем бы ни заключалась проблема, общий принцип таков: соотношение конкретики и интерпретации в системе наблюдения определяется не только тем, что желательно, но и тем, что возможно.

Таблица 4.2

**Пример микросистемы наблюдения для регистрации поведения младенца
при взаимодействии с матерью**

I. Тип вокализаций
1) отсутствуют; 2) отдельные звуки; 3) крик; 4) гуление; 5) плач; 6) хныканье; 7) смех.
II. Направление зрительного внимания
1. Направление взгляда: 1) к лицу матери; 2) от лица матери; 3) взгляд следует за лицом матери; 4) слегка в сторону на уровне носа; 5) слегка в сторону, вниз; 6) слегка в сторону, вверх; 7) в сторону на уровне носа; 8) вниз; 9) вверх.
2. Поворот головы: 1) к матери на уровне носа; 2) к матери, вниз; 3) к матери, вверх; 4) слегка в сторону на уровне носа; 5) слегка в сторону, вниз; 6) слегка в сторону, вверх; 7) в сторону на уровне носа; 8) вниз; 9) вверх.
3. Левый/правый модификатор положения головы: 1) слева от младенца; 2) справа от младенца.
4. Моргание и специфические движения глаз: 1) моргание; 2) глаза скошены; 3) взгляд в сторону и сфокусирован на видимом предмете (например, стуле), не используемом матерью в качестве элемента взаимодействия; 4) взгляд значительно смещен в сторону от оси носа.
III. Выражение лица
1. Щеки (лишь примеры): 1) нейтральное положение, расслаблены; 2) подтянуты, впалые; 3) выпуклые.
2. Брови (лишь примеры): 1) нейтральное положение покоя; 2) изогнуты, чуть приподняты посередине; 3) постоянно двигаются вверх-вниз.
3. Рот (только примеры): 1) нейтральное положение покоя; 2) слегка приоткрыт и напряжен; 3) широкая улыбка; 4) широко открыт.

4. Глаза: 1) нейтральное положение; 2) широко открыты; 3) полузакрыты; 4) закрыты.
 5. Язык: 1) не высунут; 2) чуть высунут, но не выходит за пределы губ; 3) язык высунут и выходит за пределы губ.
 6. Специфические выражения лица: 1) плачущее; 2) гримаса; 3) надутое; 4) настороженное/серьезное; 5) жмурится; 6) зевает; 7) нейтральное; 8) морщит нос; 9) спокойное; 10) простая улыбка; 11) ласковое; 12) широкая улыбка.

IV. Положение и движение тела

- 1) наклоняется вперед с согнутой спиной; 2) корпус повернут в сторону; 3) прогибается; 4) откидывается назад; 5) лежит на боку; 6) нейтральное; 7) положение изменяется матерью; 8) приподнимается; 9) туловище вытянуто, голова приподнята над подушкой или туловище и шея вытянуты; 10) наклоняется вперед с прямой спиной.

V. Движение рук и ног

1. Амплитуда движений: 1) отсутствует; 2) незначительная; 3) средняя; 4) значительная.
 2. Руки и ноги, задействованные в движении: 1) отсутствуют; 2) одна рука или одна нога; 3) две руки или ноги; 4) три конечности; 5) обе руки и обе ноги; 6) из-за матери видны только руки — двигается одна; 7) то же, что и в 6 — обе руки.
 3. Пространственная зона, в которой осуществляется движение: 1) отсутствует; 2) средняя линия; 3) между средней линией и плечами; 4) справа или слева.
 4. Специфические движения рук: 1) трет глаза; 2) тянет руки ко рту; 3) машет; 4) шевелит пальцами; 5) руки сомкнуты на уровне средней линии; 6) руки и ноги вытянуты вперед.
 5. Специфические движения ног: 1) брыкается; 2) вздрагивает.

Источник: «Analysis of Face-to-Face Interaction in Infant-Adult Dyads» (pp. 43–44) H. Als, E. Tronick & T. Berry Brazelton. In M. E. Lamb, S. J. Suomi, & G. R. Stephenson (Eds.), Social Interaction Analysis (pp. 33–76), Copyright 1979, Madison, WI: The University of Wisconsin Press.

Таблица 4.3

Пример макросистемы наблюдения для регистрации поведения младенца при взаимодействии с родителями

Дистальное/Аффилиативное поведение	
Поведение	Определение
Улыбка	Выражение лица: брови не нахмурены; уголки губ приподняты
Взгляд	Направление взгляда к родителю
Вокализации	Все типы нетрвожных вокализаций, за исключением смеха; все случаи смеха отмечаются отдельно
Предложение	Случаи, когда младенец предлагает, показывает предмет (игрушку) взрослому либо указывает на предмет
Проксимальное поведение/поведение привязанности	
Поведение	Определение
Близость	Фиксируются каждые 15 секунд нахождения младенца в радиусе метра от взрослого

Поведение	Проксимальное поведение/поведение привязанности Определение
Приближение	Движение к взрослому до расстояния 1 метр, т. е. движение в направлении «близости»
Нервозность	Любая форма вокализации, выражающая дистресс и обращенная к взрослому
Прикосновение	Фиксируется каждый случай, когда младенец касается либо тела, либо одежды взрослого
Протягивание рук	Ребенок поднимает и протягивает руку в направлении взрослого
Желание оказаться на руках у взрослого	Одно или более поведенческое проявление из следующих: ребенок нервничает, протягивает руки к взрослому, хнычет или льнет к ногам взрослого

Как наблюдать

Предположим, исследователь с ручкой и папкой в руке и твердым намерением провести наблюдение определенного рода поведения дошкольников приходит в детский сад. Как ему регистрировать интересующие его данные?

Один из возможных вариантов: просто записывать в повествовательной форме видимое поведение в его естественном виде. Этот метод называется **нарративной записью**, его также называют *записью специфических моделей* (термин предложен Баркер и Райт (Barker & Wright, 1951), *поток поведения*. Конечно, даже для максимально подробного нарративного описания все же характерна некоторая избирательность. Обычно все внимание сосредотачивается на одном ребенке, а поведение других детей учитывается лишь в момент их взаимодействия с наблюдаемым ребенком. Наблюдая ребенка, необходимо постоянно принимать решения о том, какие действия стоят того, чтобы их зафиксировать, а какие (например, моргание, сглатывание) можно проигнорировать. Кроме того, нужно принимать решения об уровне описания поведения. Например, сомкнул Джонни пальцы, сжал кулак или угрожал другому ребенку? В определенном смысле, при использовании нарративной записи наблюдатель выполняет функции видекамеры и магнитофона. Однако он — это видекамера-магнитофон со встроенным мощным аппаратом-редактором.

Несмотря на упомянутые выше ограничения, главным достоинством нарративной записи является ее подробность. Она дает больше информации, чем любой другой метод наблюдения. Эта полнота изложения делает нарративную запись особенно ценной для специалистов, которым требуется исчерпывающая информация о ребенке. Поэтому нарративную запись обычно используют учителя и клиницисты. Кроме того, нарративная запись может послужить отправной точкой для разработки исследовательских программ, наводя на мысль о существовании неких феноменов, которые затем будут изучаться более узконаправленными система-

тическими методами. Наконец, нарративная запись это не всегда и необязательно всего лишь подготовительное мероприятие перед исследованием; если регистрация производится достаточно умело и последовательно, полученные данные могут стать базой для исследования. В этих случаях нарративная запись представляет из себя «сырые» данные, которые необходимо кодировать и проанализировать; в результате из потока действий вычленяются интересующие исследователя элементы и феномены.

Если говорить о недостатках, ведение нарративной записи может потребовать значительных расходов и времени. Особенно высок уровень требований к наблюдателю, как и вероятность разных форм субъективности и систематических ошибок. У исследователя может накопиться огромный объем информации, только малая часть которой представляет некоторый интерес. Или, исследователь еще до начала наблюдения сформулировал вполне четкие цели и гипотезы, тогда нарративная запись будет не самой экономичной формой сбора данных. В обоих случаях имеет смысл использовать более узконаправленные методы наблюдения.

Второй из основных способов наблюдения — **метод временных срезов** (иногда называемый также методом *интервальных срезов*). От нарративной записи его отличают две особенности. Во-первых, при использовании метода временных срезов в центре внимания находится ограниченное количество специфических действий, а не их поток. В своем практическом применении приведенные в качестве примеров микро- и макросистема (табл. 4.2 и 4.3) также предполагают использование временных срезов. Здесь заранее оговорен точный список действий, и регистрируются только эти действия. Поскольку уже имеются четкие определения, нет необходимости составлять нарративное описание поведения; используется некий контрольный перечень или система кодирования. Второй отличительной особенностью является разделение всего периода наблюдения на строго отмеренные, обычно непродолжительные временные интервалы. Наблюдатель может 15 секунд наблюдать, отвернуться и регистрировать увиденное в течение 15 секунд, наблюдать еще 15 секунд, еще 15 секунд регистрировать и т. д. Таким образом, «срезы» в методе временных срезов происходят в двух плоскостях: анализируется только часть поведения, а само наблюдение производится только в определенные интервалы времени.

Третий способ лучше объяснить на примере. В своем, как это часто называют «классическом», исследовании Доу (Dawe, 1934) изучала ссоры дошкольников. Хотя мнения воспитателей и расходились, оказалось, что ссоры — не очень частое явление, имеющее место в среднем 3,4 раза в час. Если принять во внимание низкую частоту подобных случаев, использование нарративной записи временных срезов было бы малоэффективным. Кроме того, метод временных срезов мог бы ввести в заблуждение; наблюдатель мог бы упустить момент ссоры, если бы она произошла в период регистрации, или увидеть лишь ее часть, если бы она происходила в промежуточный момент. Поэтому Доу использовала метод **выборочного анализа поведенческих событий**, при котором единицей изучения служит эпизод изучаемого поведения, а не интервал времени. Как и при использовании метода

временных срезов, наблюдатель прежде всего должен определиться, какое поведение его интересует. Однако, используя метод выборочного анализа поведенческих событий, наблюдатель просто дожидается, пока это поведение будет иметь место, и только затем начинает записывать. Протокольные записи могут быть разных видов, от нарративного описания до последовательности кодов. Доу использовала сочетание заранее сформулированных категорий с дополнительными замечаниями. К оценивавшимся категориям относились «пассивное поведение», «мстительное поведение» и «ненаправленная активность». Какова бы ни была форма регистрации, сосредоточение на определенном виде поведения как элементарной единице наблюдения позволяет получить информацию (например, среднюю продолжительность эпизодов изучаемого поведения, предшествующие ему события и события, следующие за ним), которая могла бы быть упущена при использовании временных срезов.

В целом, факторы, влияющие на выбор той или иной системы регистрации, те же, что и факторы, влияющие на выбор микро- или макронаблюдения: цель и осуществимость. В одних обсервативных исследованиях (например, в отчетах о клинических случаях) необходимо нарративное описание, в других — уместен более узконаправленный подход с использованием таких методов, как временные срезы и выборка событий. Каковы бы ни были цели, исследователь должен применять систему, отвечающую требованиям данных конкретных условий, не превосходящую возможности наблюдателей и лимит времени, а также предполагающую рациональное соотношение между полезной информацией и затратами времени и усилий. Более подробно о плюсах и минусах разных систем регистрации можно узнать из ряда специальных работ (Hartman & Wood, 1990, Mann, Have, Plunkett, & Meisels, 1991, Odom & Ogawa, 1992).

Определение точности результатов наблюдения

Этот раздел открывает рассмотрение двух специфических факторов, способных отразиться на точности результатов наблюдения. Его завершает анализ проблемы надежности.

Поведение, которое регистрируется в обсервативном исследовании, может являться функцией от множества предшествующих и сопутствующих обстоятельств. Один из факторов, влияние которого на поведение *нежелательно*, — это само присутствие наблюдателя. Тем не менее присутствие наблюдателя и, как следствие, осознание испытуемым того, что он является объектом наблюдения, может так или иначе изменить поведение. Подобные эффекты входят в категорию *реактивности* — эффекта непреднамеренного влияния экспериментальных мероприятий на поведение испытуемого. Когда речь идет об обсервативном исследовании, эффект реактивности, как правило, называют проблемой **влияния наблюдателя**.

Вопрос о том, насколько важен эффект реактивности, уже давно является предметом научных споров. Есть основания полагать (например, Brody, Stoneman, & Wheatley, 1984; Russell, Russell, & Midwinter, 1992), что и взрослые, и дети ведут себя несколько иначе, зная, что за ними наблюдают; есть также свидетельства того,

что при определенных обстоятельствах наблюдение абсолютно не влияет на поведение. Хартманн и Вуд (Hartman & Wood, 1990) тщательно анализируют факторы, от которых зависит эффект реактивности. По их мнению, есть несколько способов снижения вероятности влияния наблюдателя. Один из них — дать испытуемым привыкнуть к присутствию наблюдателя, то есть ввести наблюдателя в обстановку до начала наблюдения и производить регистрацию только после того, как испытуемые привыкли к наблюдателю и возвратились к своему естественному поведению. Этот метод иногда называют «методом невидимки». Вариант его, который хотя и не всегда, но можно использовать, — наблюдение, производимое тем, с кем испытуемые уже хорошо знакомы, кто является естественным персонажем в данной обстановке, например родителем или учителем. Выполнение функций наблюдателя известным лицом называется **«включенным» наблюдением**.

Еще одна стратегия — сокрытие факта наблюдения. Можно, к примеру, использовать скрытую камеру или одностороннее зеркало. Конечно, такая возможность нередко отсутствует — эти методы применимы только в специальных условиях. Кроме того, скрытое наблюдение связано с этическими, а также материально-техническими ограничениями. Как мы увидим в главе 8, наблюдение за людьми без их на то согласия может рассматриваться как нарушение принципов этики.

Второй недостаток обсервативных исследований — вероятность **необъективности наблюдателя**, которая является лишь частным случаем более общей проблемы. Многочисленные исследования, начатые Робертом Розенталем (Robert Rosenthal, 1976), свидетельствуют о том, что ожидания исследователей, которые они привносят в свое исследование, иногда ведут к искажению результатов в направлении ожидаемых или желательных. Более подробно об этом рассказано в главе 5. В обсервативном исследовании присутствует риск того, что наблюдатель будет видеть и заносить в протокол только то, что ожидает увидеть, а не то, что происходит на самом деле.

Одним из доказательств может служить исследование Кент, О'Лири, Дайамент и Дитц (Kent, O'Leary, Diamant, & Dietz, 1974). Наблюдателям были продемонстрированы видеозаписи, как было сказано испытуемым, исходной фазы и фазы экспериментального воздействия некой программы, направленной на коррекцию агрессивного поведения в школе. Половине наблюдателей было сообщено, что прогнозируется снижение уровня агрессии; другой половине — что никаких изменений не предвидится. Фактически все наблюдатели смотрели одну и ту же видеозапись, в которой никаких изменений поведения не происходило. Оценивая затем эффективность программы, 9 из 10 наблюдателей, ожидавших снижения уровня агрессии, сообщали о том, что оно действительно произошло; а 7 из 10 наблюдателей, не ожидавших изменений, не отмечали никаких изменений. Интересно, что протоколы, которые вели участники обеих групп в процессе просмотра видеозаписи, были сходными, однако окончательные оценки свидетельствовали о влиянии ожидания.

Результаты этого исследования говорят о том, что для снижения вероятности необъективности наблюдателя необходимо сделать систему оценки максимально специфичной и объективной. Чем шире простор для интерпретации, тем больше у

наблюдателя возможностей исказить результаты собственной предвзятостью. Еще один способ снижения вероятности необъективности — сделать так, чтобы наблюдатель не знал о выдвинутых гипотезах или о том, к какой группе принадлежат испытуемые. Скрытие информации, которая может привести к необъективности, является условием так называемого **наблюдения «вслепую»**. Мотив его использования очевиден: если нет никаких ожиданий, нет и опасности эффекта ожидания. К сожалению, проведение наблюдения «вслепую» затруднительно, а в некоторых случаях — невозможно. Кроме того, даже если проведение наблюдения «вслепую» возможно, к нему прибегают далеко не всегда.

Еще одного рода проблемы так или иначе связаны с понятием надежности. Как отмечалось ранее, надежность означает согласованность результатов измерения. При использовании методов наблюдения ключевым моментом является единодушие наблюдателей: могут ли двое или более наблюдателей прийти к единым выводам относительно некоторого поведения? Такое согласие является необходимым условием точности результатов наблюдения. Однако выполнения этого требования недостаточно, поскольку существует вероятность того, что оба исследователя пришли к единому, но ложному выводу. Это вновь частный случай общего принципа: надежность — это необходимое, но недостаточное условие валидности.

Существует множество способов расчета надежности. Для определенного вида данных подходит коэффициент корреляции. Чем выше коэффициент корреляции между результатами, полученными двумя независимыми наблюдателями, тем выше надежность. Еще один, часто используемый индекс — процент совпадений. Предположим, оценка определенного вида поведения происходит 20 раз. Совпадение мнений двух наблюдателей в 19 случаях из 20 соответствует 93 %, достаточно высокой надежности. Совпадение только в 13 случаях из 20 соответствует 65 %, что может рассматриваться как неудовлетворительная надежность. О других методах расчета надежности, а также сложностях, связанных с каждым из них см.: Mitchell, 1979 и Hartmann, 1982.

Вопрос о том, как рассчитать надежность, встает сразу, как только исследователь осознал необходимость надежности своего исследования. При обзоре журнальных статей (см. табл. 1.3) учитывалось наличие или отсутствие рассчитанного коэффициента надежности там, где это было желательно. Результаты обзора свидетельствуют о достаточно высоком, но не полном осознании необходимости учета фактора надежности: данные о надежности были представлены в 91 % случаев из тех, где эти данные были вполне уместны. Одом и Огава (Odom & Ogawa, 1992) более подробно рассматривают расчет коэффициента надежности в обсервативных исследованиях, включая анализ типов статистических показателей и критерии их использования.

Приемы повышения надежности легко описать, но не всегда легко использовать. До начала процедуры сбора данных наблюдатели должны пройти тщательную подготовку. Система оценок должна быть максимально четкой и конкретной. Для подготовки наблюдателей и уточнений системы оценок может использоваться пилотажное исследование, в ходе которого категории редко встречающегося и трудно оцениваемого поведения либо исключаются, либо трансформируются в

более удобные категории. Наконец, если есть возможность, для того чтобы иметь непрерывную и воспроизводимую запись, лучше снимать поведение видеокамерой

Как следует из вышесказанного, при сборе данных нужно как можно раньше позаботиться о надежности. Желательно также контролировать ее на всем протяжении исследования. О целесообразности этого говорят исследования Рэйд (Reid, 1970; Taplin & Reid, 1973). В исследовании Тэплин и Рэйд наблюдатели прошли предварительную подготовку, в ходе которой достигли приемлемого уровня надежности. Затем одну группу наблюдателей проинформировали о том, что надежность оцениваться больше не будет; а участникам второй группы сообщили, что периодически, без предупреждения будет осуществляться проверка их надежности. В действительности записи всех наблюдателей продолжали сравнивать с исходным уровнем. Результат был очевиден: наблюдатели, ожидавшие, что их оценки будут подвергаться проверкам, сохраняли более высокий уровень надежности. Данная тенденция к снижению надежности изначально надежных наблюдателей при отсутствии контроля называется **дрейфом наблюдателя**. Дрейф наблюдателя относится к категории искажающих валидность факторов, которую Кэмпбелл и Стэнли (Campbell & Stanley, 1966) называют **инструментацией**: непреднамеренным изменением измерительного инструмента в ходе исследования.

До этого момента о надежности говорилось так, будто есть некий единый индекс надежности, которому исследование либо соответствует, либо нет. В действительности, в типичном исследовании, как правило, существует множество видов надежности — для определенных форм поведения, определенных аспектов поведения, определенных периодов времени, определенных подгрупп испытуемых и т. д. Суть состоит в том, что надежность должна быть продемонстрирована на том уровне, на котором производится анализ данных. Если, к примеру, исследователь хочет оценить посттестовые различия, обусловленные неким воздействием, необходимо продемонстрировать надежность посттестовых данных; надежности, достигнутой на фазе претеста, будет недостаточно. Аналогично, если изучается частота агрессивных действий, надежности общих оценок уровня агрессии будет недостаточно, исследователь должен показать, что наблюдатели сходятся во мнении и относительно частоты случаев проявления агрессии.

Ярроу и Уакслер (Yarrow & Waxler, 1979) иллюстрируют это положение интересным и несколько менее очевидным примером. Эти авторы описывают ряд консервативных исследований, в которых коэффициенты надежности подсчитывались отдельно для мальчиков и для девочек. В некоторых случаях оценка поведения представителей одного пола была более надежной, чем оценка поведения представителей противоположного пола. По крайней мере в этих исследованиях оценки агрессии мальчиков были надежнее оценок агрессии девочек. Кроме того, существенная связь индивидуальных различий в уровне агрессии с другими показателями наблюдалась только у мальчиков. Как отмечают Ярроу и Уакслер, этот результат может отражать действительно имеющиеся место различия между полами или всего лишь недостаточную надежность оценок агрессивного поведения девочек. В этом случае, как и в предыдущем, надежность необходима на уровне, на котором используются данные.

Ярроу и Уакслер (Yarrow & Waxler, 1979) также обсуждают плюсы и минусы возложения на наблюдателя функций измерительного инструмента. Их рассуждение может быть кратким выводом к этой главе:

«Даже будучи чрезвычайно опытным, наблюдатель по многим критериям — неточный научный инструмент: нестандартный, неоткалиброванный, зачастую нестабильный и ненадежный. Эти недостатки компенсируются такими присущими лишь человеку качествами, как чувствительность, гибкость и аккуратность. Задача состоит в том, чтобы решить, как провести строгое наблюдение, в полной мере используя свойственную человеку пронизательность» (р. 37).

Резюме

Эта глава начинается с рассмотрения некоторых базовых принципов измерения. Переменные, с которыми мы работаем, определяются — операционально — способами их измерения. Измерение всегда заключается в переводе некоего глобально-го конструкта (например, агрессии) в более конкретную, объективную, поддающуюся количественной оценке форму. Этот перевод подразумевает выбор из множества возможных ограниченного числа способов измерения. Необходимо выбирать аспект поведения, который будет измеряться (например, частота, интенсивность), характер объекта измерения (явный или неявный), определить, являются результаты измерения признаками или специфическими моделями изучаемого конструкта, и решить, следует сосредоточивать внимание на временном состоянии или на устойчивом свойстве.

При оценке качества измерения имеет значение следующее. Измерение должно привести к результатам, находящимся на определенном уровне и в определенных рамках, при отсутствии эффектов как *нижнего*, так и *верхнего пределов*. Если в исследование включаются разнородные группы (например, дети разных возрастов), необходимо решить вопрос эквивалентности измерения. Во избежание *монооперационального искажения* следует применять разноплановые методы. Наконец, исследователь должен позаботиться о *надежности* (согласованности) и *валидности* (точности) измерительных операций.

В следующем разделе главы, посвященном стандартизованным тестам, вопросы надежности и валидности рассматриваются более подробно. В отношении теста валидность означает, что тест измеряет именно то, что призван измерять. Анализируются три формы валидности теста: *содержательная валидность*, *критериальная валидность* и *конструктивная валидность*. Наиболее сложной из этих форм является конструктивная валидность, для установления которой могут потребоваться экспериментальные и корреляционные данные, а также обеспечение конвергентной и дивергентной валидности корреляций между показателями.

Надежность определяется как согласованность результатов измерения. Рассматриваются следующие формы надежности: *ретестовая надежность* и *надежность внутрениней согласованности*. Рассматривается также угроза для валидности, возникающая при отсутствии надежности, — *регрессия к среднему*. Регрессия — это тенденция, проявляющаяся в изменении при повторном тестировании изначально крайних показателей в направлении среднего. Вероятность регрессии особенно

велика в исследованиях с использованием экспериментального вмешательства и в определенном рода планах уравненных групп.

За разделом, посвященным тестам, следует рассмотрение другой важнейшей формы измерения — методов наблюдения. Анализируются три вопроса. Первый касается уровня специфичности, на котором оценивается поведение. При использовании *микросистемы наблюдения* учитываются мельчайшие детали поведения; *макросистема наблюдения* содержит более глобальные, интерпретационные категории. Выбор исследователем той или иной системы, находящейся на континууме от микро- к макромоделям, определяется целями исследования. Еще одной детерминантой является осуществимость: определенная система наблюдения применима только при возможности обеспечения точности результатов наблюдения. Для выбора метода регистрации также имеют значение цели и осуществимость. Наиболее подробное описание поведения дает *нарративная запись*. К более узконаправленным методам наблюдения относят *временные срезы* и *выборочный анализ поведенческих событий*. В обоих случаях категории, по которым ведется оценка, определяются заранее, наблюдение происходит в рамках либо временных блоков (временные срезы), либо определенного вида поведения (выборка событий).

Последний из рассмотренных в этой главе вопросов касается проблем, которые могут возникнуть в обсервативном исследовании. *Влияние наблюдателя* — это особая форма реактивности; она связана с тем, что люди нередко изменяют свое поведение, если знают, что являются объектом наблюдения. Обсуждается ряд методов, направленных на снижение вероятности этого искажения. Еще один из возможных источников искажения — ожидания наблюдателя, который в некоторых случаях фиксирует то, что ожидает увидеть. Наиболее эффективный способ борьбы с этим искажением — минимизация ожиданий. Наконец, надежность наблюдателей определяется совпадением их мнения при категоризации определенного поведения. Надежность следует контролировать на протяжении всего исследования, с тем чтобы предотвратить феномен *дрейфа наблюдателя*. Кроме того, она должна быть проиллюстрирована данными, соответствующими анализируемому данным.

Упражнения

В этой главе подчеркивалось, что любой теоретический конструкт может иметь множество операциональных определений. Рассмотрите следующие конструкты: альтруизм, креативность, мудрость. Для каждого дайте концептуальное определение и по крайней мере два операциональных определения. Сделайте то же самое для двух других конструктов, которые вас особенно интересуют.

Выберите некий конструкт из области психологии развития, который вас особенно интересует (например, интеллект, креативность, темперамент, Я-концепцию). Найдите по меньшей мере два стандартизованных теста, направленных на оценку индивидуальных различий по данным конструктам. Дайте критическую оценку всем методикам. Какую из них вы бы выбрали для собственного исследования и почему?

Выполнение следующего задания предполагает, что у вас есть возможность привлечь к исследованию по крайней мере одного родителя с ребенком в возрасте около года. Достаньте статью Лэм (Lamb, 1976), о которой говорилось в тексте, изучите обсервативную систему оценки и попытайтесь применить ее на практике. Имейте в виду, что вам придется адаптировать лабораторную процедуру к домашним условиям; однако по крайней мере некоторые элементы системы Лэм вполне воспроизводимы в домашней обстановке. *Упражнение будет особенно полезным*, если вы попытаетесь провести исследование с разными парами родитель—ребенок. Кроме того, хорошо было бы объединиться с кем-то из сокурсников и рассчитать коэффициент надежности ваших наблюдений.

Глава 5

Процедура

До этого момента мы рассматривали вопросы, связанные с планированием исследования. Как отобрать испытуемых, каким экспериментальным воздействиям их подвергнуть? Как лучше всего сравнивать разные возрастные группы и экспериментальные условия? Как, иными словами, составить план исследования, которое даст валидные результаты?

Несомненно, разрешение этих вопросов является необходимым условием грамотного проведения исследования; однако, как уже говорилось в главе 1, этого недостаточно. Разрабатывая план и систему оценок, мы лишь подходим к началу исследования; принятые решения нужно воплотить в жизнь, тестируя или наблюдая испытуемых. При реализации задуманного может возникнуть масса проблем. Именно этим проблемам — тому, что может привести к ошибочным результатам, — и посвящена эта глава. Или, другими словами, в этой главе говорится о том, как избежать проблем и получить валидные результаты.

Следует еще раз подчеркнуть то, на что мы обращали внимание в главе 1. Знание общих принципов проведения исследования не может избавить от всех проблем. В конечном счете, эти принципы предстоит применить на практике, в процессе живого общения с испытуемыми и использования конкретных методик.

Стандартизация

Я начну с рассмотрения того, что, несомненно, является необходимым условием для проведения большинства исследований — стандартизации. **Стандартизация** означает создание единообразия всех аспектов процедуры для всех испытуемых. Для проведения стандартизации исследователь должен заранее решить, как именно будет реализовываться каждый аспект процедуры — как будет сформулирована инструкция, как и когда будет предъявляться стимульный материал и т. д. Приняв решения о стандартизации, исследователь должен на практике проконтролировать их выполнение.

Основная причина проведения стандартизации обсуждалась в главе 2 — необходимость экспериментального контроля. Исследователь должен иметь возможность контролировать ряд аспектов экспериментальной обстановки: точную фор-

му независимых переменных, способ измерения зависимых переменных, а также другие факторы ситуации. При отсутствии такого контроля невозможно ни узнать наверняка, что именно изучается в данном исследовании, ни проинтерпретировать полученные результаты.

Рассмотрим некоторые примеры затруднений, которые могут возникнуть при отсутствии адекватной стандартизации. Допустим, проводится исследование кратковременной памяти у школьников, скажем, у 1- и 4-классников. Экспериментатор показывает по одной серии картинок с изображением знакомых предметов, и дети получают задание вспомнить и назвать как можно больше из увиденного. В табл. 5.1 представлены разнообразные варианты отклонения экспериментатора от стандартной процедуры и, следовательно, искажения результатов. Экспериментатор может, к примеру, непреднамеренно варьировать время экспозиции для разных картинок или для разных детей; или забыть сказать некоторым детям, что порядок воспроизведения свободный, что достаточно существенно при попытке вспомнить последовательность, состоящую из множества предметов; или же экспериментатор может дать некоторым детям больше попыток, повышая их шансы на успех.

Таблица 5.1

Варианты отклонений от стандартной процедуры и их эффекты при изучении памяти

Аспект процедуры	Предполагаемая процедура	Отклонение от стандартной процедуры	Возможные эффекты
Формулировка инструкций	Сообщение всем детям о том, что воспроизведение возможно в любом порядке	Только части детей сообщается о том, что порядок воспроизведения не имеет значения	Общие: появляется незапланированная дисперсия, и результаты уже не поддаются однозначной интерпретации Специфические и искажающие: если формулировка инструкции неодинакова для разных возрастных групп, возрастные различия могут искажаться или искусственно создаваться
Время экспозиции	Все дети видят картинки в одинаковое время	Время экспозиции неодинаково для разных картинок или для разных детей	Общие: те же, что и в предыдущем случае Специфические и искажающие: если время экспозиции неодинаково для разных возрастных групп или для разных экспериментальных условий, тогда различия между группами могут получить неправильное истолкование
Количество попыток	Все дети получают равное количество попыток	Одни дети получают больше попыток, чем другие	Общие: те же, что и в предыдущем случае Специфические и искажающие: если дети из разных возрастных групп получают разное количество попыток, возрастные различия могут искажаться или искусственно создаваться

Как явствует из табл. 5.1, такие погрешности в стандартной процедуре могут вызвать разного рода проблемы. В лучшем случае, недостаточная стандартизация приведет к неопределенности, поскольку будет неясно, в чем состоят экспериментальные условия, и поэтому непонятно, как интерпретировать полученные результаты. Неопределенность может исказить и дескриптивную информацию, которую дает стандартизованное исследование: как узнать, к примеру, насколько ответ определенного испытуемого выделяется на общем фоне, если мы не знаем точного времени экспозиции и количества попыток? Она может отразиться и на внешней валидности результатов. Если то, что обнаруживается, зависит от некоего (незапланированного и неизвестного) сочетания условий, полученные результаты нельзя перенести на тот контекст, на который исследователь намеревается их перенести.

Недостаточная стандартизация может исказить и внутреннюю валидность в том случае, если отклонения от стандарта систематически распределены между разными группами неравномерно. Предположим, что экспериментатор дает старшим детям больше попыток, чем младшим (возможно, из убеждения, что правильный ответ скорее придет на ум старшим детям). В этом случае выявленные возрастные различия могут в действительности отражать изменения в процедуре. Мы придем к ложному выводу о причинно-следственных связях, от чего пострадает внутренняя валидность.

Таким образом, стандартизация имеет огромное значение. Может показаться, что провести стандартизацию довольно просто — все, что необходимо, это заблаговременно составить план процедуры, а затем действовать в соответствии с этим планом. Но на практике, как известно каждому, кто имеет исследовательский опыт, все не так просто. Трудно предусмотреть все проблемы, которые могут возникнуть в ходе исследования, трудно также строго придерживаться плана долгой и сложной процедуры при непосредственном взаимодействии с испытуемыми. Эти трудности усугубляются еще и тем, что в большинстве исследований с детьми общению с ребенком желательно придать неформальный, естественный характер, что является лучшим способом создания у ребенка ощущения комфорта и поддержания его интереса. Жесткое следование плану — самый легкий путь стандартизации — может нарушить естественность, которая необходима для поддержания положительного настроения и активности ребенка.

Вероятно, основным условием оптимальной стандартизации является наличие опыта, как проведения исследования вообще, так и работы с определенным возрастным контингентом, использования определенных приемов. Хотя исследовательское мастерство приходит со временем, приобретение необходимого специфического опыта возможно в процессе *пилотажного исследования* — экспериментирования с исследовательскими приемами и приобретения навыка их использования до начала самого исследования. Тщательное пилотажное исследование дает возможность заблаговременно устранить дефекты и сомнительные моменты и, таким образом, иметь на руках схему стандартизованной процедуры еще до появления перед экспериментатором первого испытуемого. Кроме того, в ходе пилотажного исследования экспериментатор учится уверенной, непринужденной манере инструктирования испытуемых и взаимодействия с ними, учится преодолевать скованность

и избавляется от привычки работать по бумажке. В процессе пилотажного исследования экспериментатор учится превосходить и быть готовым прореагировать на некоторые проблематичные ситуации, когда поведение испытуемого вынуждает отступить от стандартной процедуры. Часто в исследовании с детьми наиболее важным качеством экспериментатора является не столько умение следовать стандартной схеме, сколько мастерство индивидуального подхода к каждому ребенку, при котором результатам эксперимента не грозит искажение, а ребенок не чувствует отчуждения.

Отклонения от стандартной процедуры

Стандартизация желательна, но не является безусловной целью. Нередко имеет смысл удовлетвориться неполной стандартизацией. Мы только что рассмотрели подобный случай: ситуацию, в которой поведение ребенка вынуждает к отступлению от стандартной процедуры. В этом разделе мы проанализируем еще несколько примеров.

Начнем с того, что вспомним замечание, сделанное в главе 2: полная стандартизация, в смысле создания абсолютно равных условий для всех испытуемых, невозможна. Если, к примеру, тестирование проводится индивидуально, нельзя протестировать всех испытуемых в 10 часов утра 30 сентября; время тестирования неизбежно будет разным у разных испытуемых. Невозможно поддерживать на одном уровне и такие факторы, как расписание занятий в день тестирования, то, насколько хорош обед, погоду и ряд других обстоятельств, влияющих на результаты каждого ребенка. Как говорилось в главе 2, важно не то, чтобы все испытуемые находились в одинаковых условиях, а то, чтобы любые потенциально значимые факторы были равномерно распределены между сравниваемыми группами. При этом мы получаем возможность исключить смешение любого фактора с интересующими нас независимыми переменными.

В других случаях вынужденное отклонение от стандарта обусловлено попыткой объединить в одном исследовании разные возрастные группы. Рассмотрим исследование, направленное на оценку успешности решения задач испытуемыми в возрасте от дошкольного до окончания колледжа (например, Zelniker, Oppenheimer, & Renan, 1975). Одной из независимых переменных здесь является возраст. Действуя в соответствии с данной выше рекомендацией, мы должны были бы избегать возможных изменений в процедуре при изучении разных возрастных групп. Однако маловероятно, что исследователь будет вести себя с 5-летним испытуемым так же, как с 20-летним. Условия, в которых будет проводиться исследование, скорее всего будут различаться: детей будут тестировать в одном из школьных кабинетов, а взрослых — в университетской лаборатории. Инструкции, по крайней мере частично, тоже будут различаться: 20-летнего студента не будут приглашать «пойти поиграть». Даже если слова инструкций одни и те же, изменяется темп и тон речи; естественная манера разговора неодинакова для 5- и 20-летних. Наконец, в процедуру могут быть внесены и существенные изменения; маленьким детям иногда необходима помощь. Например, исследование Зелникера и др. (Zelnicker et al., 1975) включало предварительную тренировку только для самых младших.

Сказанное выше не означает, что если между испытуемыми большая разница в возрасте, стандартизация и контроль не нужны. Необходимость стандартизации в рамках одной возрастной группы сохраняется. Кроме того, центральные моменты процедуры, например условия задач в исследовании Зелникер и др., должны быть неизменными для всех испытуемых; иначе не имеет смысла проводить возрастное сравнение. Однако, что касается других аспектов процедуры, при изучении разных возрастных групп исследователь должен попытаться достичь не буквальной, а функциональной эквивалентности процедуры. Иными словами, цель — сделать так, чтобы процедура была *равно соответствующей возрасту* для всех изучаемых возрастных групп. Условия школьной библиотеки отличаются от условий университетской лаборатории; однако их обстановка одинаково естественна и знакома для изучаемых групп. В этой ситуации уравниваются не характеристики помещений самих по себе, а уровень естественности обстановки при тестировании. Тот же аргумент можно предложить и для обоснования адаптации к определенному возрасту слов инструкции, тона речи, обратной связи, похвалы и т. д. Каждый раз исследователь пытается проводить процедуру в соответствии с возрастом испытуемого, и в то же время, насколько это возможно, сохранять ее единообразие для разных возрастных групп. Трудность здесь заключается в том, чтобы, адаптируя процедуру для каждой из возрастных групп, не исказить результаты их сравнения.

Последний вариант отклонения от стандарта — в **эксплоративном исследовании**. Как следует из названия, эксплоративное исследование — это исследование, цель которого в разработке новой, малоизученной проблематики оригинальными, не определенными заранее способами. Суть такого исследования — в возможности творческого подхода, в возможности разобраться в неожиданных результатах, опробовать различные методы изучения, поэкспериментировать с самой процедурой, модифицировать ее в ходе проведения исследования. Классическими примерами эксплоративного исследования являются многие из работ Пиаже (например, Piaget, 1926; Piaget & Szeminska, 1952). Как станет понятно из дальнейшего рассуждения, важнейшим элементом успеха Пиаже была его способность открывать и анализировать новые феномены, придерживаясь гибкого стиля при постановке вопросов и формулировании задач, руководствуясь как ответами ребенка, так и нормами процедуры. Именно такого рода экспериментирование и изменения исключаются в жестко стандартизированном исследовании.

В дальнейшем, говоря о Пиаже, мы отметим еще один момент. Для оценки валидности выводов Пиаже потребовалось провести множество более строго контролируемых стандартизованных исследований. Эксплоративное исследование, как никакое другое, подходит для открытия новых феноменов и генерирования идей, однако верификация этих феноменов и идей зависит от точности и контроля, которые могут обеспечить только стандартизованные процедуры.

Сверхстандартизация

Как мы только что увидели, стандартизация не может быть абсолютной. Полная стандартизация, без преувеличения, невозможна. По мере исследования испытуемых что-то всегда изменяется. Несмотря на все старания исследователя стандартизировать процедуру, поведение испытуемого может вынудить к отступлению от

установленной схемы. Изучая разные возрастные группы, исследователь намеренно изменяет определенные параметры процедуры. Исследователь, разрабатывающий новую проблематику, намеренно отказывается от стандартизации ради возможности свободного экспериментирования и совершения открытий.

Но предположим, что мы имеем дело с ситуацией, не соответствующей ни одному из вышеперечисленных исключительных случаев. Допустим, мы работаем с испытуемыми одного возраста, имея четкий план действий, контролируя ситуацию, не изучая нечто новое, а пытаюсь с максимальной точностью определить, каким образом связаны некие переменные. Есть ли основания в этом случае строго придерживаться стандарта? Или существует, даже здесь, такое понятие, как излишняя стандартизация?

Кэмпбелл и Стэнли (Campbell & Stanley, 1966) полагают, что сверхстандартизация действительно бывает. Приведенный ими пример касается исследования способности к убеждению, в котором сравнивается эффективность обращения к рассудку и эмоциям. Чтобы произвести стандартизацию, исследователь решает записать на магнитофон по одной версии обращения обоих типов; таким образом, всем испытуемым в определенной экспериментальной ситуации предъявляется один и тот же стимул. Такой контроль кажется желательным, поскольку мы должны были бы знать наверняка, что представляют из себя определенные экспериментальные условия. Однако в действительности решение сохранять ситуацию неизменной означает отказ от возможности выяснить, какие именно аспекты экспериментальных условий объясняют обнаруженный эффект. Возможно, имеет значение тот фактор, который мы намеревались изучить: обращение к эмоциям или к рассудку. Но возможно, что имеет значение пол говорящего, или некоторые особенности его голоса, тон или темп его речи. Возможно также, имеет значение определенное сочетание вышеперечисленных качеств — к примеру, апеллирование к эмоциям колоритным мужским голосом. Суть в том, что любой фактор, одинаковый для всех испытуемых, потенциально становится частью независимой переменной. Предпочтительнее в такой ситуации допустить разнообразие этих дополнительных факторов (пола, особенностей голоса и т. д.), в то же время оставляя неизменным фактор, который нас действительно интересует: характер обращения. Тогда любой обнаруженный эффект можно с большей уверенностью объяснить характером обращения.

Понятно, что сверхстандартизация может быть причиной искажения внешней валидности. В описанном выше исследовании цель — получить в той или иной мере обобщенные выводы относительно эффективности апелляции к рассудку или эмоциям. Для нас нежелательно, чтобы результаты зависели от индивидуальных качеств говорящего или от способа подачи текста, так как полученные в этом случае выводы нельзя считать справедливыми для конкретного исследования.

Трудности, создаваемые сверхстандартизацией, доказывают значение, которое в научном исследовании имеет *воспроизведение*. При воспроизведении, по определению, сохраняются неизменными те характеристики, на изучение эффекта которых направлено само исследование; к примеру, обращение к эмоциям или к рассудку в гипотетическом исследовании эффективности убеждения. В то же время другие характеристики практически неизбежно изменяются, особенно если иссле-

дование воспроизводит другой исследователь, в другой лаборатории. Поэтому человек, выступающий в роли убеждающего, скорее всего будет другим, как и экспериментатор, взаимодействующий с испытуемыми, а также помещение, в котором проводится эксперимент, время года и т. д. Обнаружив те же эффекты эмоционального и рационального убеждения, несмотря на все изменения, мы сможем более уверенно говорить о том, что тот или иной результат действительно обусловлен характером обращения. И наоборот, если бы нам не удалось воспроизвести изначально полученные результаты, это означало бы, что параметры процедуры, считавшиеся несущественными, в действительности не так уж несущественны.

Когда сверхстандартизация становится проблемой в исследовании психологии развития? Пожалуй, чаще всего в случаях, аналогичных тем, которые описали Кэмпбелл и Стэнли: когда все данные собирает один экспериментатор. В обзоре журналов по психологии развития (табл. 1.3) учитывалось в том числе и количество экспериментаторов, принимавших участие в исследованиях, в которых могли иметь значение характеристики экспериментатора. В 30 % отчетов из тех, в которых имелась подобная информация, указывалось на наличие одного экспериментатора; только в 22 % случаев количество экспериментаторов достигало четырех. Обратите внимание, что по меньшей мере 69 % статей при данном подсчете не учитывалось, поскольку в них не содержалась информация о количестве экспериментаторов; в действительности в 52 % отчетов вообще отсутствовала какая-либо информация об экспериментаторах.

Неудивительно, что исследования, проводимые одним экспериментатором столь распространены. Использование одного экспериментатора требует меньше расходов, меньше времени на подготовку и не связано с проблемой составления графика проведения исследования и подбора экспериментаторов. Каким бы заманчивым ни казался этот выбор, всегда помните о том, что если экспериментатор работает в одиночку, то становится константным элементом экспериментальной процедуры, элементом, который может влиять на результаты исследования. Совпадение результатов, полученных разными экспериментаторами, служит залогом внешней валидности исследования.

Некоторые систематические ошибки

Мы пришли к заключению, что недостаточная или чрезмерная стандартизация может создать определенные проблемы. В этом разделе будет рассмотрено еще несколько, теперь уже более специфических причин искажения валидности. При этом я ссылаюсь на работы Кэмпбелл и Стэнли (Campbell & Stanley, 1966), а также Кук и Кэмпбелл (Cook & Campbell, 1979), выводы которых сведены в табл. 2.4. Мы проанализируем четыре специфические причины: инструментарий, систематическую ошибку при отборе испытуемых, фактор истории и реактивность.

Как отмечалось в главе 4, *инструментарий* — это фактор изменения физических инструментов измерения либо экспериментаторов и наблюдателей в ходе проведения исследования. Поэтому он попадает в категорию неудовлетворительной стандартизации: некоторый аспект процедуры, который должен быть неизменным, в действительности изменяется в процессе исследования. Хотя в некоторых слу-

чаях изменения затрагивают инструментарий в буквальном смысле (например, секундомер, неисправно работающий в сырую погоду), чаще всего изменения касаются человека, выполняющего роль экспериментатора или наблюдателя. К примеру, в ходе исследования экспериментатор приобретает мастерство и уверенность при инструктировании испытуемых. Или наоборот, процесс работы утомляет и разочаровывает экспериментатора, и он начинает допускать небрежности. В любом случае, процедура, которую проходят первые испытуемые, будет отличаться от той, которую проходят следующие.

Наибольшее значение проблема инструментария приобретает тогда, когда изменения, происходящие в ходе исследования, смешиваются с одной из изучаемых независимых переменных. Рассмотрим исследование, в котором осуществляется сравнение успешности выполнения 1-классниками и 2-классниками некоей экспериментальной задачи. Исследователь решает протестировать большинство или всех 1-классников прежде, чем тестировать 2-классников. Для этого может быть множество причин: возможность работать только с одним учителем в одном кабинете, получение разрешения родителей только в одном классе, «нетерпение пойти поиграть» детей из одной группы, желание учителей, чтобы тестирование в их классе прошло как можно быстрее, и т. д. Однако исследователь, уступающий давлению факторов удобства, рискует упустить из вида потенциально значимый источник смещения. Если есть хоть какая-то вероятность изменения экспериментатора в ходе проведения исследования, различия между 1- и 2-классниками нельзя будет интерпретировать однозначно.

Вопрос инструментария встает не только при сравнении разных возрастных групп, но и при сравнении двух или более экспериментальных условий. Часто в исследовании с несколькими условиями легче всего сначала собрать данные для одного условия, а затем — для другого, поскольку в этом случае в одном сеансе задействуется только один набор материалов и методик. Однако вновь цена удобства — возможное смешение. Лучше, если бы время, когда проводится тестирование в одних экспериментальных условиях, примерно соответствовало времени, когда проводится тестирование в других экспериментальных условиях.

Есть и еще один источник систематических ошибок в ситуации, когда переход к другому экспериментальному условию происходит только после завершения работы с предыдущим. Это систематические ошибки, обусловленные повсеместным требованием получения родительского разрешения на проведение практически любого исследования с детьми. Допустим, исследователь разослал по домам детей одного из классов бланки разрешений. 15 бланков вернулись незамедлительно, и эти 15 детей тестируются в условиях 1. Неделю спустя были возвращены еще 15 бланков, и эти дети тестируются в условиях 2. В этом случае валидность может исказить (помимо фактора инструментария) *систематическая ошибка при отборе*. Вполне вероятно, что родители, заполнившие бланки незамедлительно, отличаются от тех, кто затягивал процедуру. Тогда и дети из обеих групп, скорее всего, будут различаться независимо от изменения условия 1 на условие 2.

Необходимость предотвращения смешения факторов возраста или экспериментальных условий со временем тестирования подтверждается еще одним фактом. Как бы ни были постоянны поведение экспериментатора и процедуры, в ходе ис-

следования могут изменяться другие потенциально значимые факторы. Это хорошо известно любому исследователю, который хоть раз пытался протестировать школьников накануне праздника или просто днем в пятницу. Исследование — это лишь одно из множества событий в жизни испытуемых, и необходимо позаботиться о том, чтобы это множество событий не привело к смешению при каком бы то ни было сравнении.

От сказанного выше легко перейти к рассмотрению переменной, которая в работах Кэмпбелл и Стэнли (1966) названа историей. В том смысле, в котором этот термин употребляют Кэмпбелл и Стэнли, *история* связана с вопросом, который мы только что затронули: с влиянием внешних по отношению к исследованию событий на результат этого исследования. Особое значение данная проблема приобретает в планах «претест-посттест», в которых некое постороннее событие ведет к появлению альтернативного объяснения обнаруженных изменений. Допустим, мы изучаем отсроченное удовлетворение желания, то есть способность детей отказываться от немедленного получения небольшой награды ради получения через некоторое время чего-то более привлекательного (исследование Паттерсон и Картер (Patterson & Carter, 1979), кратко описанное в главе 2, направлено на изучение именно этого понятия). В нашем исследовании ребенок стоит перед выбором: получить одну конфету сразу или подождать до следующего дня и получить несколько конфет. Оценив способность детей к отсрочиванию удовольствия, мы вводим экспериментальное воздействие, направленное на развитие этой способности. Экспериментальная программа длится всю вторую половину октября, а повторное тестирование осуществляется в начале ноября. Мы обнаруживаем, что после прохождения экспериментальной программы детям легче устоять перед соблазном немедленного получения сладостей. Является ли это подтверждением эффективности наших действий? Необязательно, если учесть, что на один из дней между первой и второй проверкой выпадает Хэллоуин, а следовательно, и море сладостей. Имеет место смешение экспериментального воздействия и фактора истории, которое затрудняет интерпретацию результатов.

О факторе истории уже говорилось в главе 3. Тогда в качестве примера приводилось исследование эффективности некой образовательной программы, направленной на повышение успеваемости неблагополучных детей. При этом было найдено решение проблемы, связанной с фактором истории: использование контрольной группы, не подвергающейся экспериментальному воздействию, — то есть группы, переживающей те же жизненные события, что и экспериментальная группа, за исключением экспериментального воздействия. В случае исследования способности отсрочивать удовлетворение желания мы могли бы просто протестировать какую-то группу детей до и после обильных праздничных угощений, обычных для Хэллоуина¹. Однако заметьте, что использование контрольной группы не решает всех проблем, обусловленных влиянием внеэкспериментальных событий. В определенных случаях эти события взаимосвязаны с экспериментальным воздействием. Возможно, что естественно, получаемый опыт празднования Хэллоуина и опыт, полученный при прохождении составленной нами программы, эф-

¹ Конечно, более легким решением было бы перенесение исследования на другое время.

фективны только в сочетании друг с другом: ни угощения, ни экспериментальное воздействие сами по себе не влияют на способность ребенка отсрочивать удовлетворение желаний. В этом случае использование контрольной группы не прояснит ситуацию. Тогда мы столкнемся с серьезным ограничением на генерализацию полученных результатов.

Реактивность и связанные с ней проблемы

Следующая причина искажения валидности заслуживает того, чтобы посвятить ей отдельный раздел. Как указывалось в табл. 2.4, термин *реактивность* означает непреднамеренное воздействие экспериментальных мероприятий на поведение испытуемых — или, проще говоря, тот факт, что, находясь под наблюдением, люди ведут себя не так, как в естественной обстановке. В случае со взрослыми эффект реактивности обусловлен осознанием испытуемыми того, что они принимают участие в эксперименте. Дети же чаще всего не осознают, что являются объектом исследования; в действительности самые маленькие из них вообще этого не понимают. Тем не менее выполнение роли объекта исследования может изменить поведение кого угодно, и поэтому реактивность может стать проблемой при изучении любых возрастных групп.

Исследование со взрослыми испытуемыми связано с рядом специфических проблем, относящихся к категории реактивности (Silverman, 1977). Особого упоминания заслуживают две из них. Суть первой проблемы легче всего объяснить на примере. Мартин Орн (Martin Orn, 1962) попытался найти задание, которое можно было бы использовать для изучения гипнотического контроля, — задание, которое было бы столь скучным, что все не впавшие в гипнотическое состояние испытуемые отказывались бы продолжать его выполнение. Одним из опробованных заданий было решение примеров на сложение случайных чисел. Каждый листок содержал по 224 примера, а каждому испытуемому выдавалась пачка из 5000 листов. Пять с половиной часов спустя некоторые испытуемые все еще были заняты работой! Тогда Орн добавил еще одно условие: покончив с каждым листом, испытуемый должен был разрывать его на 32 кусочка. Несмотря на явную бессмысленность этого задания, некоторые испытуемые продолжали работу, пока их в конечном счете не остановил экспериментатор.

Это исследование, как и другие исследования Орна, демонстрирует тот факт, что люди, принимающие участие в эксперименте, иногда (хотя, конечно, не всегда — критические замечания по этому поводу можно найти у Берковитц и Доннерстайн (Berkowitz & Donnerstein, 1982)) способны пойти на многое, если считают, что это необходимо экспериментатору. Поведение «идеального испытуемого» может принять форму и общей готовности следовать указаниям, и отдельных попыток подтвердить любую гипотезу, которая, по мнению испытуемых, положена в основу исследования. Используя термин Орна, можно сказать, что испытуемые реагируют на характеристики запроса экспериментатора, то есть «на совокупность признаков, по которым испытуемый судит об экспериментальной гипотезе» (Orne, 1962, p. 779). Такая тенденциозность реакций испытуемого приводит к тому, что называется *эффектом выдвигаемых требований*.

Вторая причина систематических ошибок — «эффект ожидания оценки» (Rosenberg, 1965); еще один термин, используемый для сравнения с «идеальным испытуемым», — «самолюбивый испытуемый» (Silverman, 1977). Здесь речь идет о том, что испытуемые своим поведением стараются заслужить высшие оценки экспериментатора; иными словами, пытаются произвести благоприятное впечатление. Разумеется, создание благоприятного впечатления по смыслу может совпадать с выполнением того, что хочет экспериментатор; в этом случае фактор выдвигаемых требований и фактор оценки имеют общий эффект. Однако они не всегда синонимичны; например, испытуемый может попытаться удивить экспериментатора своими «высокоинтеллектуальными» ответами. Силвермен (Silverman, 1977) приводит данные, свидетельствующие о том, что в борьбе фактора запроса и фактора оценки выигрывает, как правило, последний.

Острее всего проблемы, связанные с ожиданием оценки, стоят в исследованиях с использованием **самоотчетов**. Как следует из названия, при использовании метода самоотчета экспериментальные данные представляют из себя вербальные отчеты людей о самих себе — о своих качествах, своем опыте, типичном поведении и т. д. В таком исследовании для испытуемого очень легко и заманчиво подкорректировать ответы в сторону желаемого, а не описывать то, что есть на самом деле. Например, в исследовании, направленном на изучение методов воспитания, мать может утверждать, что она никогда не бьет ребенка, хотя в действительности время от времени прибегает к физическому наказанию.

Описанные в этом разделе эффекты исследовались в основном на взрослых испытуемых. Насколько велико их значение в детской психологии? Очевидно, что при изучении маленьких детей нам нет необходимости беспокоиться об эффекте запроса и об эффекте оценки в том смысле, в каком они описаны для взрослых. В действительности, исследователь, изучающий младенцев или 2–3-летних детей, может даже сожалеть, что его маленькие испытуемые столь безразличны к целям исследования и к стремлениям самого исследователя. Установку на сотрудничество можно приветствовать, одновременно помня об опасности эффекта запроса. Непривычность и неестественность экспериментальной ситуации может повлиять даже на самых маленьких детей. Вполне можно ожидать, что при обследовании в лаборатории младенец будет проявлять большую тревогу и беспокойство, чем при обследовании дома. При общении с незнакомым взрослым дошкольник может замкнуться или же, увлекшись стимульным материалом, обрадовавшись вниманию взрослого, говорить о чем угодно и делать все, что угодно, кроме того, что нужно исследователю. А к школьному возрасту у детей начинают появляться некоторые из классических признаков реактивности, которые наблюдаются у взрослых. Любому, кто хотя бы однажды тестировал школьников, легко вспомнить ребенка, который по лицу взрослого пытается понять, что требуется сделать; который повышает голос в конце каждого ответа, будто спрашивая: «Верно?»; который в целом более всего обеспокоен тем, чтобы угодить взрослому, тем, чтобы произвести приятное впечатление, или и тем и другим.

Как минимизировать эффект реактивности? Поскольку реактивность обусловлена осознанием себя объектом изучения, очевидным решением является сокрытие того факта, что идет исследование. Степень сокрытия может быть разной.

В простейшем случае исследователи представляют маленьким детям свои экспериментальные задания как игры, а не «тесты» или «эксперименты». Такое объяснение зачастую достаточно точно, более понятно для ребенка, чем сложные термины, и скорее всего не вызовет такой тревоги, как приглашение «протестироваться». Еще один характерный прием при изучении дошкольников — игра исследователя со своими маленькими испытуемыми, целью которой является установление доверительных отношений до начала самого эксперимента. Разумеется, нужно помнить о том, что использование этих приемов снятия напряженности не гарантирует отсутствия тревоги или сопротивления; ребенок все равно понимает, что происходит нечто необычное. Кроме того, экспериментатор должен следить, чтобы предупредительные учителя не делали напрасными его усилия, приглашая ребенка «пойти протестироваться у доктора Такого-то».

Вероятность реактивности минимальна в исследованиях, проводимых в обстановке, максимально приближенной к естественной. Рассмотрим исследование, направленное на изучение того, какие игрушки предпочитают дети дошкольного возраста. Исследователь может выполнить эту задачу, приводя каждого ребенка в комнату, где проходит эксперимент, и там предлагая специальный опросник или тест (называя их при этом «игрой», а не «тестом»). В этом случае измерение является непосредственным и эффективным, однако вероятность реактивности (тревоги в связи с необходимостью отвечать на вопросы, угадать, что хочет услышать взрослый, и т. п.) крайне высока. Еще один способ — также привести ребенка в комнату, где проводится эксперимент, но теперь уже просто наблюдать, с какой из игрушек, находящихся в комнате, он захочет играть. Если предложение поиграть будет звучать естественно и непринужденно (например, «Ты можешь поиграть, пока я закончу свою работу»), вероятность эффекта реактивности будет невелика. Третий способ — наблюдать детей в естественной обстановке игровой комнаты детского сада; в конечном счете, дошкольники значительную часть времени проводят за игрой. Если исследователю удастся понаблюдать за детьми, оставаясь для них невидимым (к примеру, через одностороннее зеркало), тогда реактивность вообще не будет проблемой. Наконец, в некоторых случаях о поведении можно судить по его физическим эффектам, без наблюдения за самими испытуемыми. Популярность разных игрушек можно, к примеру, определить, отметив в конце дня, какие из игрушек стоят нетронутыми на полках, а какие — разбросаны по комнате. Если проводить более длительное изучение, о популярности игрушек можно судить по тому, насколько они истрепаны; какие из игрушек все еще как новенькие к концу года, а какие потерты и сломаны? Здесь, конечно, возможность реактивности полностью отсутствует. В книге Уэбб, Кэмпбелл, Шварц, Секрест и Гроу (Webb, Campbell, Schwartz, Sechrest, & Grow, 1981) подробно обсуждается использование таких приемов «невмешательства» при изучении моделей поведения. В книге Лагрека (LaGrec, 1990) можно найти ряд рекомендаций по снижению эффектов реактивности в ситуации, когда для сбора данных необходима личная беседа с испытуемым.

Установка на определенный ответ

Обратимся теперь к проблеме, тесно связанной с проблемой реактивности. **Установка на определенный ответ** — это склонность испытуемого отвечать на вопрос или выполнять задание неким заранее определенным способом, независимым от

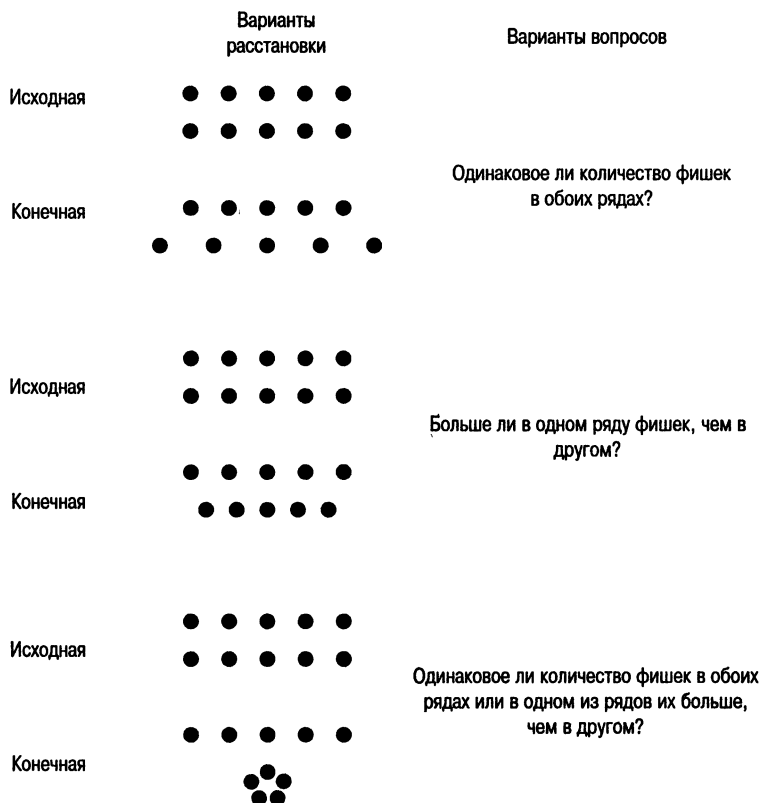


Рис. 5.1. Варианты расстановки предметов и вопросов при изучении представлений о сохранении числа

содержания задания. Согласно этому определению, поведение «идеального испытуемого», которое рассматривалось как проявление реактивности, можно считать одним из видов установки на определенный ответ: испытуемый стремится сказать или сделать то, что, как ему кажется, ждет от него экспериментатор, а не выполняет задание само по себе. В этом разделе мы проанализируем установки на определенный ответ, которые в большей мере зависят от испытуемого, а не от экспериментальных условий. Однако нужно признать, что граница между описанными выше проблемами реактивности и установками на определенный ответ, которые рассмотрены ниже, довольно прозрачна и условна. В обоих случаях мы имеем дело с искаженными ответами, которые могут привести к невалидным выводам.

Принципы, изложенные в этом разделе, легче будет понять на конкретном примере. Рисунок 5.1 иллюстрирует задание, которое более подробно будет рассматриваться в главе 11: задача Пиаже на сохранение числа (Piaget & Szeminska, 1952). Если говорить о числе, сохранение означает осознание того, что количество предметов в наборе не изменяется исключительно из-за их перестановки. На рис. 5.1 изображены несколько вариантов подобных трансформаций, которые можно использовать

для выявления наличия или отсутствия представлений о сохранении. Рисунок также иллюстрирует разные способы формулировки вопроса о сохранении.

Какие установки на ответ могут возникнуть у детей при решении задачи на сохранение? Типичная форма установки — «*соглашательство*»: склонность отвечать утвердительно на любой вопрос. Очевидно, что оно становится потенциальной проблемой всегда, когда задается однонаправленный вопрос, как в первых двух примерах на рис. 5.1. Дети, которых каждый раз спрашивают: «Одинаковое?» и которые каждый раз отвечают «да», могут в действительности понимать принцип сохранения, хотя их ответы и не имеют никакого отношения к стоящей перед ними задаче.

Простое соглашательство можно предотвратить использованием двунаправленных вопросов, как в третьем примере на рисунке. Однако возникают и другие проблемы. Некоторые дети склонны выбирать *вариант, названный последним*, то есть соглашаться со всем, что стоит на последнем месте в вопросе взрослого. Если постоянно задавать вопрос так, как в третьем примере, результатом будет вывод об отсутствии (но вполне вероятно, псевдоотсутствии) представлений о сохранении. Другие дети склонны *чередовать ответы*, давая то один ответ, то другой, в зависимости лишь от того, в который раз им задан вопрос. При изучении сохранения чередование ответов может наблюдаться и в рамках одной задачи, и от одной задачи к другой, поскольку вопрос о количестве, как правило, задается и до и после изменения расстановки. Наконец, при выполнении некоторых заданий, включая и задачи на сохранение, определенные трудности могут создавать позиционные предпочтения. Маленький ребенок, к примеру, может постоянно отвечать, что больше всего предметов в ближайшем к нему ряду.

Следует сделать несколько общих замечаний, касающихся установок на определенный ответ. Во-первых, хотя я и привел в качестве примера задание на сохранение, возможность подобного затруднения ни в коем случае не ограничена случаями изучения сохранения или, коли на то пошло, изучения дошкольников. Всегда, когда испытуемый должен дать вербальный ответ, присутствует вероятность предпочтения утвердительного ответа. И, как будет понятно из главы 10, определенные виды позиционных предпочтений наблюдаются уже у новорожденных.

Второе замечание касается интерпретации установок на определенный ответ. Что означает склонность ребенка отвечать «да» на любой вопрос о сохранении? Эту тенденцию вполне можно принять за свидетельство непонимания принципа сохранения — ребенок, прибегающий к столь элементарной форме ответа, вряд ли имеет представление о феномене сохранения. И действительно, такая интерпретация зачастую оправдана. Проблема в том, что это нельзя выяснить наверняка. Постоянно указывая на более длинный ряд, как на содержащий большее количество предметов, ребенок дает однозначно неправильный ответ. Ребенок же, который отвечает только утвердительно, просто не выполняет задание — возможно, не понимая изучаемого феномена, возможно, сбивтый с толку формулировкой вопроса, возможно, не имея достаточной мотивации к тщательному обдумыванию или по любой другой причине. Установки на определенный ответ не дают сделать никаких вы-

водов, за исключением вывода о том, что у ребенка есть установка на определенный ответ.

Это подводит нас к последнему замечанию: исследователь должен стремиться минимизировать вероятность установок. Существуют разные подходы к решению этой задачи. В случае с сохранением исследователь должен использовать простой и понятный язык, предложить ребенку потренироваться до начала тестирования, мотивировать ребенка к обдумыванию ответов и т. п. Но в некоторых, даже самых изощренных исследованиях установки на определенный ответ неизбежны. В такой ситуации необходимо, чтобы исследователь по крайней мере осознал их наличие, иначе он неправильно истолкует поведение испытуемого. Исследователь, который проверяет понимание принципа сохранения, предлагая лишь одно задание, никогда не сможет с уверенностью сказать, что означает ответ ребенка. Получение ответов на ряд вопросов, сформулированных по-разному, является более надежным основанием для вывода о том, понимает ли ребенок принцип сохранения и проявляет ли в своих ответах установку.

Коммуникация между испытуемыми и диффузия

В эксперименте исследователей Хорка и Фэрроу (Horka & Farrow, 1970) была обнаружена довольно странная систематическая ошибка в ответах, не соответствующая ни одной из описанных в предыдущем разделе категорий. В их исследовании от испытуемых (5- и 6-классников) требовалось идентифицировать ряд букв, выделяющихся на фоне черных бессмысленных фигур. Используемый стимульный материал изображен на рис. 5.2. Половине детей, тестируемых утром в один день, предъявляли стимул, изображенный внизу рисунка; а тестируемым днем того же дня — стимул, изображенный сверху. На рассмотрение рисунка отводилось 4 минуты, а за правильный ответ ребенок получал в награду 50 центов.

Искажение ответов наблюдалось у испытуемых, проходивших тестирование днем. Многие из них сообщали, что видят слово *left* («левый»), то есть давали ответ, который был бы верным утром. Частота этого ответа приблизительно в два раза превышала частоту ответа, правильного для дневного тестирования. Кроме того, днем ответ *left* давался в два раза чаще, чем это было утром, когда *left* было именно тем словом, которое показывали испытуемым!

Причиной этому, конечно, являлось то, что некоторые из испытуемых, проходивших тестирование утром, общались с некоторыми из испытуемых, которых должны были тестировать днем. По окончании процедуры всем детям, тестируемым утром, сообщили правильный ответ и напомнили, что они смогли бы получить 50 центов, всего лишь сказав слово *left*. Их также просили никому не рассказывать об исследовании. По всей видимости, это предупреждение не смогло заставить детей отказаться от шанса помочь другу или просто проявить свою осведомленность о необычном событии.

Исследование Хорка и Фэрроу доказывает, что причиной искажения ответов испытуемых бывает любая информация об исследовании, которую они получают от других испытуемых. Коммуникация между испытуемыми ведет к разнообразным последствиям. Иногда испытуемые успешнее справляются с заданием, потому что другие испытуемые подсказали верный ответ или обратили внимание на

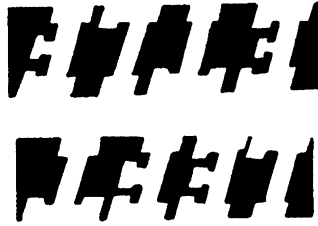


Рис. 5.2. Стимульный материал, использовавшийся Хорка и Фэрроу при изучении коммуникации между испытуемыми. (S. Horka & B. Farrow, 1970, *Journal of Experimental Child Psychology*, 10, p. 364)

некий аспект процедуры, который должен был держаться в секрете. Иногда испытуемые хуже справляются с заданием, потому что процедура отличается от той, которую проходили другие испытуемые (как в исследовании Хорка и Фэрроу), или потому что услышанный ими рассказ об исследовании слишком сбивчив, чтобы служить подспорьем. Любой, кто хоть однажды слышал, как маленькие дети рассказывают о своих впечатлениях от исследования, знает, насколько далеко бывает их описание от того, что действительно происходило во время эксперимента. Эффект коммуникации может быть и специфичным, проявляясь в виде повышения или снижения интереса, настороженности и т. д., в зависимости от того, что слышали испытуемые. Каковы бы ни были конкретные эффекты, их общим следствием является искажение, которое может сделать невалидными полученные результаты.

Насколько общий характер носит проблема коммуникации испытуемых? Очевидно, что есть такие исследования, в которых эта проблема не стоит — к примеру, исследования младенцев (если не помешают родители!), исследования, в которых испытуемые либо не знакомы друг с другом, либо не имеют возможности общения в процессе исследования. Если взять противоположную ситуацию, исследование Хорка и Фэрроу имеет ряд особенностей, максимизирующих вероятность коммуникации. В их число входит простота, легкая вербализация правильного ответа и награда за него, что характерно далеко не для всех исследований. По этой причине нужно быть осторожными при распространении результатов, полученных Хорка и Фэрроу, на все исследования. Тем не менее, как отмечают Хорка и Фэрроу, их процедура действительно имеет много общего с исследованиями, в которых дети обучаются или решают задачи, имеющие правильные ответы, которые могут достаточно свободно передаваться одним ребенком другому. Кроме того, изучение детей довольно часто происходит в школьной обстановке, при этом испытуемые хорошо знают друг друга и имеют практически неограниченные возможности для общения. И действительно, из всех областей психологического исследования именно исследование со школьниками чаще всего связано с проблемой коммуникации испытуемых. Поэтому со стороны исследователя, работающего в школе, будет разумно принять меры по предотвращению подобной коммуникации, а также использовать методы, направленные на выявления ее эффектов. В своей работе Брукс и Кендалл (Brooks & Kendall, 1982) анализируют способы минимизации коммуникации между испытуемыми.

Один из эффектов коммуникации настолько значим, что ему дали собственное название. Под термином **диффузия** подразумевается незапланированное распро-

странение эффекта экспериментального воздействия с экспериментальной группы на контрольную, не подвергающуюся воздействию. Рассмотрим в качестве примера программу вмешательства в дошкольном возрасте, направленную на улучшение подготовленности к школьному обучению детей, входящих в группу риска неуспеваемости (например, Gray, Ramsey, & Klaus, 1982). Типичным подходом в таких исследованиях является разделение испытуемых на две группы: экспериментальную, подвергающуюся воздействию определенной программы, и равноценную ей экспериментальную группу, не подвергающуюся воздействию этой программы. Однако, если обе группы детей проживают в одном квартале, на протяжении периода исследования наверняка будут контакты и коммуникация: дети будут играть друг с другом, а родители — беседовать и обмениваться впечатлениями. Поэтому эта программа может благотворно подействовать не только на экспериментальную группу; дети из экспериментальной группы, которые по замыслу исследователя не должны были подвергаться экспериментальному воздействию, также могут получить определенный опыт и продемонстрировать некоторое улучшение. Именно в связи с риском размывания границ между экспериментальной и контрольной группами при оценке эффективности программ вмешательства используются две контрольные группы: *проксимальная контрольная группа*, которую набирают из района проживания экспериментальных испытуемых, и *дистальная контрольная группа*, которую набирают из другого района. Использование последней, «бесконтактной» группы позволяет оценить вероятность диффузии.

Предубеждения экспериментатора

Мы обсудили уже ряд систематических ошибок, включая те, которые допускает исследователь, совершая просчеты при проведении исследования. За исключением, возможно, лишь эффекта выдвигаемых требований, рассмотренные выше типы искажения по сути являются ненаправленными, в том смысле, что они могут с равной вероятностью работать как на пользу исследованию, так и ему во вред. Обратимся теперь к тому, что, по-видимому, можно считать самой коварной формой искажения: к систематическому искажению результатов исследования, связанному с субъективизмом самого исследователя. Такое систематическое искажение приводит к так называемому **эффекту предубеждений экспериментатора**; другое, часто употребляемое название — *эффект исследовательских ожиданий*. О вероятности подобного эффекта упоминалось в главе 4 при обсуждении обсервативного исследования; в этом разделе мы рассмотрим эту проблему в более общем виде.

Первым, кто занялся вопросом предубеждений экспериментатора, был Роберт Розенталь (Robert Rosenthal, 1976). В типичном исследовании Розенталя испытуемым, чаще всего студентам, предлагается выступить в роли экспериментаторов. Этих экспериментаторов случайно делят на две группы. В качестве элемента подготовки каждой группе сообщают о том, что ожидает (и, возможно, надеется) обнаружить ведущий исследователь. Группы получают абсолютно противоположные сведения об ожидаемых результатах. Одной группе, к примеру, могут сказать, что ожидаются высокие результаты, а другой — что ожидаются низкие результаты. Не считая этой разницы в ожиданиях, обе группы получают одинаковую подготовку

и следуют единому плану проведения эксперимента. Несмотря на это сходство, группы нередко получают разные результаты, и это различие отражает характер ожиданий каждой из групп. Иными словами, экспериментаторы обнаруживают то, что ожидают обнаружить. Подобные эффекты были продемонстрированы при изучении разнообразной проблематики и различных возрастных групп, включая и детей. На 1994 год насчитывалось 464 исследования, посвященных эффекту ожидания (Rosenthal, 1994a).

Каким образом оказывают воздействие ожидания исследователя? Барбер и Силвер (Barber & Silver, 1968) описали 11 ситуаций искажения экспериментатором результатов исследования. Причины искажения перечислены в табл. 5.2. Наибольший интерес из них представляют первые пять: те, где искажение опосредовано непреднамеренными, зачастую практически незаметными действиями. Есть данные, полученные Розенталем и другими исследователями, что любой из 11 видов искажения может сыграть свою роль. Однако нужно отметить, что определить, где конкретно происходит искажение, подчас бывает невозможно.

Таблица 5.2

Причины эффекта необъективности экспериментатора

Непреднамеренные действия
1. Экспериментатор может влиять на поведение испытуемого, непреднамеренно подавая паралингвистические сигналы — например, изменяя тон речи.
2. Экспериментатор может влиять на поведение испытуемого, непреднамеренно подавая кинетические сигналы — например, изменяя позу или выражение лица.
3. Экспериментатор может влиять на поведение испытуемого непреднамеренным вербальным подкреплением.
4. Экспериментатор может непреднамеренно неверно истолковать поведение испытуемого.
5. Экспериментатор может непреднамеренно неверно зарегистрировать поведение испытуемого.
Преднамеренные действия
6–10. Экспериментатор может намеренно совершить описанные выше действия.
11. Экспериментатор может сфабриковать экспериментальные данные.

Источник: Т. Х. Barber & М. J. Silver. *Psychological Bulletin Monograph Supplement*, 1968, 70, 1–29.

Исследования Розенталя имеют и другие недостатки (анализ этого вопроса можно найти у Barber, 1976; Barber & Silver, 1968; Rosenthal, 1968). Возможно, важнейшим из них является недостаточный уровень внешней валидности. Во многих исследованиях, демонстрирующих эффект необъективности, в роли экспериментаторов выступают совершенно неопытные студенты, набранные исключительно для исследования. Каждому экспериментатору недвусмысленно дается понять, каковы предполагаемые результаты. В некоторых случаях экспериментаторам обещают повысить оплату, если полученные результаты подтвердят гипотезу исследователя. Иными словами, задействуются все механизмы максимизации эффекта ожидания. Однако все эти исследования могут просто доказывать тот факт, что

существует вероятность неграмотного проведения исследования. Самый общий вывод, который можно сделать на основе этих результатов (то есть уровень внешней валидности), заключается в том, что исследования в целом имеют некоторые погрешности, которые проявляются при изучении эффекта ожидания.

Зная о причинах искажения, можно предложить способы сведения их к минимуму. Если участники исследования, которые будут проводить тестирование, недостаточно опытные, они, естественно, должны пройти тщательную подготовку. В этой подготовке следует акцентировать внимание на проблеме стандартизации, так как отклонение от стандарта — это первый шаг к искажению. Если есть возможность, экспериментаторов нужно не только обучить проведению стандартной процедуры, но и периодически *проверять* их работу, дабы удостовериться в том, что в ходе исследования она не стала небрежной. Нет необходимости говорить, что повышение оплаты за получение желательных результатов неприемлемо. Кроме того, исследователь должен избегать дифференцированного использования *неденежного* вознаграждения — например, выражение одобрения, когда экспериментатор сообщает о желательных результатах, и неудовольствия — при отсутствии таковых. Поскольку эффект необъективности обусловлен ожиданиями, нельзя допускать, чтобы у экспериментаторов сформировались четкие представления о том, каковы должны быть результаты исследования. Иногда с этой целью экспериментаторов держат в неведении относительно гипотез, проверяемых исследованием, и статуса испытуемых (например, экспериментальная или контрольная группа). Это метод «слепого» эксперимента, который рассматривался в главе 4.

Хотя предложить способы предотвращения необъективности экспериментатора довольно просто, воплотить их на практике не всегда легко. В частности, иногда бывает трудно устроить «слепой» эксперимент или «слепое» наблюдение. В некоторых случаях ведущий исследователь (то есть человек, разработавший план исследования) тоже тестирует испытуемых, что, естественно, исключает возможность эксперимента, «слепого» в отношении гипотез. В других случаях было бы несерьезно надеяться на то, что у экспериментаторов, даже не обладающих никакой информацией об исследовании, не сформируются гипотезы о его целях и возможных результатах. Во многих ситуациях от экспериментатора невозможно скрыть статус испытуемого. Экспериментатор, к примеру, не спутает 3-летнего ребенка с 5-летним или мальчика с девочкой, а знание возраста или пола испытуемого может вызвать предвзятое отношение. Стоит отметить, что исследователи, работающие в области психологии развития, фактически всегда обладают информацией о своей главной независимой переменной — о возрасте испытуемого. Наконец, даже тогда, когда проведение «слепого» тестирования возможно, оно часто не используется. При обзоре журналов (таблица 1.3) учитывалось, использовались ли «слепой» эксперимент или «слепое» наблюдение там, где это было необходимо. Только 63 % исследователей сообщили об использовании этого приема.

Ранее мы поставили под сомнение внешнюю валидность исследований необъективности экспериментатора. Суть вышесказанного сводится к тому, что результаты этих исследований не стоит слишком быстро сбрасывать со счетов как ограниченные искусственностью созданной экспериментальной ситуации. Несомненно,

эффекты необъективности экспериментатора существуют. Как часто они имеют место и сколько ложных выводов явилось их следствием — спорный вопрос. Цель исследователя — предупредить подобные споры в отношении своего исследования, то есть при планировании и проведении исследования, а также при сообщении его результатов сводить к минимуму вероятность необъективности экспериментатора.

Выбывание испытуемых

Большая часть этой главы была посвящена проблемам, которые могут возникнуть при работе с испытуемыми. В некоторых случаях эти проблемы столь серьезны — а также столь очевидны, — что сохранить данные какого-то испытуемого не представляется возможным. Выбывание или отсев испытуемых может происходить по ряду причин, обсуждавшихся в этой главе, — из-за неприемлемого отступления от стандартной процедуры со стороны экспериментатора, из-за крайне высокой реактивности со стороны испытуемого, из-за жесткой, непреодолимой установки испытуемого на определенный ответ. Кроме того, некоторые группы испытуемых и некоторые виды исследования могут создавать особые проблемы, приводящие к выбыванию даже при идеальном проведении исследования. Маленький ребенок может уснуть или расплакаться в ходе проведения эксперимента. При проведении лонгитюдного исследования испытуемые могут переехать в другую местность или умереть.

Главное, что следует запомнить об отсевах испытуемых, это то, что его нужно свести к минимуму. Отсев может вызвать ряд затруднений. С практической точки зрения, он связан с напрасной потерей времени и экспериментатора, и испытуемого. Если количество выбывших существенно, исследователь может остаться с выборкой, слишком маленькой для того, чтобы делать какие либо выводы. Наконец — и самое главное, — избирательное выбывание ставит под сомнение валидность исследования. Эта проблема рассматривалась в главе 3. Уход испытуемых действительно часто бывает избирательным и характерен для наименее компетентных, наименее мотивированных, наименее склонных к сотрудничеству с незнакомыми людьми и т. п. Такое выбывание может отразиться на внешней валидности или возможности генерализации результатов, поскольку исследователь будет изучать совсем не ту популяцию, на которую он хочет перенести свои выводы. Если выбывание зависит от возраста испытуемых или от экспериментальных условий, оно может отразиться и на внутренней валидности.

Методы минимизации выбывания можно вывести из описанных в этой главе проблем и соответствующих рекомендаций. Здесь стоит напомнить два общих принципа. Во-первых, до начала тестирования детей, особенно дошкольного возраста, с ними нужно установить доверительные отношения. Во-вторых, до начала осуществления исследования желательно провести пилотажные проверки в том объеме, в котором это необходимо. Исследователь, методики которого не подходят для большинства испытуемых, вероятно, просто проигнорировал фазу пилотажной проверки.

Помимо сведения к минимуму количества выбывших у исследователя есть еще две задачи. Во-первых, он должен заранее определить максимально объективный

критерий отказа от испытуемого. С самого начала исследователю нужно иметь четкие представления о том, какое поведение испытуемого делает невозможным использование полученных результатов. Конкретные критерии будут разными в разных исследованиях; при изучении младенцев это может быть сон или плач; при изучении дошкольников — неудачное выполнение вербального претеста; при изучении старших детей и взрослых — распознавание испытуемым обманного приема экспериментатора. Суть в том, что если сначала собрать все данные испытуемого, проверить, согласуются ли они с выдвинутой гипотезой, и только затем побеспокоиться о критериях выбраковки данных, можно получить искаженные результаты. Опасность заключается в том, что исследователь может легко поддаться соблазну сохранить результаты, согласующиеся с его ожиданиями, и исключить те результаты, которые с ними не согласуются.

О второй задаче мы уже упоминали в главе 2: предоставление точных данных о критериях выбраковки данных испытуемых и о количестве выбывших испытуемых. Как было сказано в главе 2, эта задача не всегда выполняется: подобная информация содержалась только в 50 % журнальных статей. Эта цифра говорит о некотором прогрессе, учитывая, что в обзоре, представленном в первом издании этой книги, данный показатель составлял 36 %. Тем не менее она оставляет желать лучшего.

Следует отметить еще один момент. Я постоянно подчеркиваю необходимость сохранения как можно большего количества испытуемых. Это, однако, не означает, что рыдающего младенца или перепуганного дошкольника нужно принудить продолжать участие в эксперименте. Против подобного образа действий можно выдвинуть два аргумента. Во-первых, результаты таких испытуемых вряд ли будут иметь какую-либо ценность. Во-вторых, что более существенно, как мы увидим в главе 8, важнейшим этическим принципом исследования, проводимого на людях, является право испытуемого в любой момент выйти из эксперимента. Этим правом, как любой другой испытуемый, обладает и не умеющий говорить младенец. А права испытуемых всегда стоят выше желаний экспериментатора.

Резюме

В этой главе рассматривается воплощение абстрактного исследовательского плана в конкретную процедуру. В ней анализируются проблемы, которые могут возникнуть при работе с испытуемыми, а также способы их разрешения.

Центральным понятием данной главы является понятие *стандартизации* — создания условий, при которых важнейшие элементы экспериментальной процедуры будут одинаковыми для всех испытуемых. Стандартизация — это аналог контроля в экспериментальном исследовательском плане. При отсутствии стандартизации процедура изменяется от испытуемого к испытуемому, возникают систематические ошибки, а результаты не поддаются однозначной интерпретации. Возможно, наилучший способ стандартизации — тщательная проработка всех практических моментов на этапе *пилотажного исследования*. Лучшим является тот экспериментатор, который способен быть естественным, следуя стандартам, а также, когда это необходимо, найти подход к каждому испытуемому.

Хотя проведение стандартизации желательно, некоторое отступление от стандартной процедуры неизбежно и зачастую является разумным шагом. Полная стандартизация невозможна, поскольку некоторые аспекты процедуры (например, точное время тестирования) от испытуемого к испытуемому будут изменяться. При изучении разных возрастных групп может возникнуть необходимость внести в процедуру некоторые коррективы, чтобы она *соответствовала возрастным особенностям* разных групп. Цель *эксплоративного исследования* — открытие интересных феноменов в новой области исследования, и поэтому гибкость здесь имеет большее значение, чем стандартизация. Наконец, любой неизменный аспект процедуры (например, экспериментатор, осуществляющий сбор данных) потенциально становится частью независимой переменной, и подобная *сверхстандартизация* может ограничить возможность генерализации полученных результатов.

Затем речь идет о некоторых конкретных факторах, негативно отражающихся на валидности. Фактор инструментария подразумевает незапланированное изменение либо физических инструментов, либо экспериментаторов и наблюдателей в ходе исследования. Это изменение приобретает особое значение, если его ошибочно принимают за результат влияния возрастных особенностей или экспериментальных условий, что может произойти в ситуации, когда все или большинство испытуемых из одной группы проходят тестирование до начала тестирования другой группы. Еще одна причина, по которой следует избегать смешения условий и порядка тестирования, — *вероятность систематической ошибки при отборе*: испытуемые, с готовностью соглашающиеся на участие в исследовании (или родители, быстро соглашающиеся на участие своих детей), могут определенным образом отличаться от тех, кто затягивает с ответом. Третье основание для контроля порядка тестирования — возможность эффекта внешних по отношению к исследованию событий, например, эмоционального возбуждения накануне праздника. В планах «претест-посттест» те же события могут явиться фактором искажения валидности, который называют фактором *истории* — изменением, обусловленным действием неконтролируемых событий, происходящих в процессе исследования.

Очень часто искажение валидности связано с *реактивностью* — искажающим влиянием выполнения роли объекта исследования на поведение испытуемого. Лучше других изучены две формы реактивности: поведение «идеального испытуемого», когда человек стремится вести себя так, как того желает экспериментатор; и поведение «самолюбивого испытуемого», когда человек стремится произвести благоприятное впечатление. Чаще всего для снижения вероятности реактивности используются методы, направленные на сведение к минимуму очевидности манипуляций и измерений. Тесно связана с предыдущей проблема *установки на определенный ответ* — склонность испытуемого отвечать заранее определенным образом, независимым от содержания задания. К установкам на определенный ответ, характерным для детей, относят «соглашательство», выбор варианта, названного последним, чередование ответов и *позиционное предпочтение*. Исследователь должен также исключить возможность *коммуникации между испытуемыми и диффузии*, при которых искажение обусловлено общением испытуемых.

Особенно коварной является ситуация, когда то, что обнаруживает исследователь, зависит от его предположений и ожиданий. Подобные *эффекты предубеж-*

дений экспериментатора принимают разнообразные формы. Хотя вопрос распространенности этого феномена является спорным, возможность необъективности экспериментатора существует и потому необходимо принимать меры по ее предупреждению. Если есть возможность, нужно проводить «слепое» тестирование, с тем чтобы исключить искажающий эффект ожиданий. Важны также предварительная подготовка и периодическая проверка экспериментаторов, гарантирующие стандартное проведение процедуры исследования.

В некоторых случаях проблемы, обсуждавшиеся в главе, столь серьезны, что от испытуемого приходится отказываться. Рассматривается несколько моментов, связанных с отсевом испытуемых. Во-первых, отсев следует свести к минимуму, так как он может повлиять и на внутреннюю и на внешнюю валидность исследования. Во-вторых, исследователь должен заранее определить критерий отказа от испытуемого и отразить в своем отчете, сколько испытуемых выбыло и по каким причинам. И наконец, испытуемых нельзя принуждать к продолжению участия в исследовании против их воли.

Упражнения

Упражнения к последующим главам этой книги содержат рекомендации по практической работе, то есть варианты простых исследований, которые вы сможете провести с разными группами испытуемых. Выполнение этих заданий предполагает, что у вас есть возможность осуществления некоторых из этих исследований; лучше всего, если вы выберете достаточно сложную процедуру (не просто раздачу опросников) и сможете опробовать ее в разных ситуациях. Следите за своими успехами как экспериментатора, приобретая опыт инструктирования, предъявления стимульного материала и проведения исследования в целом. Возможно, вам захочется записать процедуру эксперимента на магнитофон для последующего прослушивания; полезным может оказаться и получение обратной связи от испытуемых. Учитывайте следующие моменты: строгое следование стандартному плану процедуры, естественность поведения, ясность инструкций, исключение возможности эффектов запроса или предубеждений экспериментатора.

В этой главе упоминалась книга Уэбба «Unobtrusive Measures» (Webb et al., 1981) и приводился пример применения данного подхода в психологии развития. Придумайте как можно больше собственных примеров. Если есть возможность, достаньте экземпляр книги Уэбба и сравните собственные идеи с рассматриваемыми в этой книге методами и примерами.

Условия проведения исследования и его цели

В этой главе рассматриваются две большие темы. Первая касается обстановки, в которой проводится исследование. Большая часть того, что уже было сказано по этому поводу, отвечало на вопрос «как?» — как грамотно спланировать исследование, как провести его с минимумом проблем и систематических ошибок. Но психолога, более чем остальных ученых, должен волновать вопрос «где?», от ответа на который, в свою очередь, может зависеть ответ на вопрос «как?». Психологическое исследование в отличие от исследований в других научных областях можно проводить в разной обстановке, начиная от жестко контролируемой искусственно созданной среды лаборатории и кончая повседневными естественными условиями детского сада, игровой площадки или супермаркета. Каждый вид условий имеет свои плюсы и минусы, а также определенную степень пригодности для изучения конкретной проблематики, что, естественно, и является причиной использования разных видов условий. Эти плюсы и минусы — первая большая тема этой главы.

Вторая большая тема касается целей, на достижение которых направлены исследования в области психологии развития. Список целей потенциально может содержать столько пунктов, сколько существует исследований, имеющих самостоятельные цели. Однако в общем цели исследования делятся на две большие категории, о которых мы говорили в главе 1, обсуждая мотивы проведения исследований в области психологии развития. Мы проводим исследования в интересах *фундаментальной науки* — для развития знаний о развитии человека; а также в *прикладных целях* — для улучшения жизни детей и других незащищенных групп населения. Две эти задачи отнюдь не противоречат одна другой, не изменяются и базовые методологические принципы при переходе от фундаментальных к прикладным вопросам.

Тем не менее прикладное исследование связано с особыми трудностями, которые накладываются на трудности исследования вообще. Поэтому в разделе, посвященном целям, рассматриваются в основном разнообразные формы прикладных исследований в области психологии развития.

Условия проведения исследования

Существуют разные способы классификации условий проведения исследования. Система, которую я использую для первоначальной классификации, взята из статьи Росса Парке (Ross Parke, 1979) и в общем виде представлена в табл. 6.1. Первое разграничение, которое вводит Парке — это разграничение на полевые и лабораторные условия. Полевое исследование осуществляется в естественной для испытуемого среде — к примеру, в упоминавшихся ранее условиях игровой площадки или супермаркета. Лабораторное исследование проводится в специально созданных условиях, в условиях, которые могут резко отличаться от естественной для испытуемого среды, и в которые испытуемого ставят исключительно в целях исследования. Передвижная исследовательская лаборатория, размещенная в трейлере, где измеряется изменение частоты сердечных сокращений в ответ на определенные звуки, подаваемые в наушники, — наглядный пример лабораторных условий.

Таблица 6.1

Классификация условий проведения исследования

		Локус зависимой переменной	
		Лабораторные условия	Полевые условия
Локус независимой переменной	Лабораторные условия	1	2
	Полевые условия	3	4

Источник: «International Designs», R. D. Parke. In R. B. Cairns (Ed.), *The Analysis of Social Interactions: Methods, Issues, and Illustrations* (p. 15–35), 1979, Hillsdale, NJ: Lawrence Erlbaum Associates.

В приведенных выше примерах граница между полевой и лабораторной обстановкой кажется довольно четкой. Однако часто она гораздо неопределеннее. Лабораторные условия могут в той или иной мере приближаться к условиям естественной среды, а полевые условия могут тут же потерять свою естественность при введении экспериментального контроля и процедур измерения. Поэтому различие между полевыми и лабораторными условиями правильнее рассматривать на континууме, а не как дихотомию. Кроме того, «естественность» — это не монолитный конструкт; существует ряд параметров, по которым может изменяться степень естественности условий. Парке рассматривает три таких параметра: неспецифическая физическая среда, стимульное поле и социальные агенты, присутствующие в ситуации. В связи с этой комплексностью дальнейшее использование терминов «полевые условия» и «лабораторные условия» следует рассматривать как упрощение, удобное при обсуждении методологии, однако искажающее более сложную реальность.

Второй фактор в системе классификации Парке имеет отношение к разграничению зависимой и независимой переменной. Независимой переменной можно манипулировать либо в лабораторных, либо в полевых условиях. Зависимую переменную можно измерять либо в лабораторных, либо в полевых условиях. Сочетание независимых и зависимых переменных с лабораторными и полевыми условиями дает четыре варианта исследовательского плана, представленных в клетках табл. 6.1.

Использование конкретного примера поможет сделать более понятным дальнейшее изложение. Рассмотрим уже знакомую нам проблематику: связь между насилием на телеэкране и агрессией. Приводит ли просмотр детьми агрессивных телепрограмм к повышению агрессии? Это интересный, тщательно изученный, но остающийся спорным вопрос. Поскольку его можно исследовать — и он действительно исследуется — с использованием каждого из четырех подходов, представленных в табл. 6.1, для наших целей он послужит подходящим примером. В дальнейшем, как в этой главе, так и в последующих, мы встретим и другие примеры. В действительности, основная проблема состоит в том, чтобы определить, какого рода вопросы можно изучать с применением каждого из этих четырех подходов.

План 1: «лаборатория—лаборатория»

Исследования, проводимые по этому плану, вероятно, более других соответствуют представлениям большинства людей о том, что такое «психологический эксперимент». Экспериментальные манипуляции производятся в контролируемых лабораторных условиях, эффект этих манипуляций оценивается в той же лаборатории. В случае изучения связи насилия на телеэкране и агрессии типичным является следующий порядок работы. Испытуемых произвольно делят на две группы: экспериментальную группу, которая будет смотреть телевизионные ролики с эпизодами насилия, и контрольную, которая будет смотреть телепрограммы нейтрального содержания. Детей из обеих групп по одному приводят в экспериментальную комнату и там показывают телесюжет, предназначенный для их группы. Чуть позже ребенку предоставляется возможность проявить агрессию в тех же лабораторных условиях. Более высокий уровень агрессии детей из экспериментальной группы по сравнению с уровнем агрессии детей из контрольной группы может рассматриваться как свидетельство роли телевидения в формировании определенной модели поведения.

В рамках этой базовой парадигмы возможен ряд вариаций. В качестве примеров могут служить исследования Бандуры, Росс и Росс (Bandura, Ross, & Ross, 1963a), а также Либерт и Бэрона (Liebert & Baron, 1972) (последнее мы уже рассматривали в главе 4, говоря о принципах измерения). Телесюжеты, которые показывают детям, могут в большей или меньшей степени походить на те, что обычно транслируются по телевидению. В одних случаях их специально монтируют для исследования (Bandura et al., 1963a), в других — показывают эпизоды коммерческой телепродукции (Liebert & Baron, 1972). Агрессивные действия также отличаются по степени их приближенности к реальному агрессивному поведению. Иногда в качестве объекта агрессии выступает неодушевленный предмет (в исследовании Бандуры и его коллег — это кукла Бобо); а иногда — другой ребенок (Liebert & Baron, 1972). В одних случаях агрессивное поведение физически идентично реальным проявлениям агрессии (удары, пинки в исследовании Бандуры); в других — оно отличается от реального (нажатие на кнопку в исследовании Либерт и Бэрона). Наконец, агрессивные реакции различаются по тому, насколько легко их измерить. Реакция нажатия на кнопку в исследовании Либерт и Бэрона может фиксироваться автоматически; а для учета ударов и пинков обычно требуется на-

блюдатель, который выносит решения о наличии или отсутствии определенных действий и об их значении.

Исследование в контролируемых лабораторных условиях имеет два больших преимущества. Первое — это контроль независимой переменной. Изучая эффекты просмотра агрессивных телепрограмм, исследователь может выбирать, какие именно телесюжеты показывать детям, какие именно дети будут смотреть определенный телесюжет и каков именно будет контекст просмотра телесюжетов. Как уже было сказано, для получения четких выводов о причинно-следственных связях необходим как раз такого рода контроль. В итоге в исследовании типа «лаборатория—лаборатория» создаются наилучшие условия для максимизации внутренней валидности.

Второе преимущество лабораторного исследования связано с зависимой переменной. По определению, зависимая переменная может свободно изменять свое значение и поэтому никогда не бывает под контролем исследователя. Однако зависимую переменную нужно измерить, и это измерение легче всего произвести в структурированной обстановке лаборатории. В лаборатории, к примеру, можно заснять на видеопленку агрессивное поведение, а затем использовать ее для воспроизведения и анализа — роскошь, зачастую недоступная на игровой площадке детского сада или в домашней обстановке. Можно также отказаться от услуг наблюдателя и работать с результатами по большей части автоматического регистрирования. Нажатие на кнопку в исследовании Либерт и Бэрона — пример автоматического фиксирования агрессивных действий; специальная кукла Бобо, фиксирующая каждый получаемый удар (Deur & Parke, 1970), — еще один пример. Наконец, как мы вскоре увидим, некоторые зависимые переменные просто нельзя измерить нигде, кроме как в контролируемых лабораторных условиях.

Лабораторный подход имеет не только достоинства, но и недостатки. Если все достоинства можно объединить в категорию «контроль», то недостатки можно определить словом «искусственность». Лабораторная обстановка может в разной степени приближаться к реальным жизненным условиям, которые нас интересуют; однако она всегда в той или иной мере отличается от них, и подчас весьма значительно. Поэтому встает вопрос: можно ли переносить результаты, полученные в лаборатории, на естественную среду? Это проблема компромиссного выбора между внутренней и внешней валидностью, о которой говорилось в главе 2. Как мы тогда отметили, факторы, повышающие внутреннюю валидность, зачастую снижают внешнюю.

Рассмотрим проблему связи просмотра телепрограмм с эпизодами насилия и агрессии. Несмотря на то что в предоставлении возможности посмотреть телевизор нет ничего необычного, ситуация просмотра в лаборатории неизбежно отличается от ситуации просмотра дома. Начать с того, что ребенок находится вне дома, где он чаще всего смотрит телевизор. Кроме того, иная не только физическая, но и социальная обстановка. Обычно ребенок смотрит телевизор в компании братьев, сестер или друзей; теперь же он смотрит телевизор в одиночестве. Еще одно отличие — взрослый явно предлагает ребенку посмотреть некий конкретный телесюжет. Показ телематериалов взрослым человеком может расцениваться ребенком как одобрение того, что происходит на экране, и как разрешение на подражание,

чего не бывает в домашней обстановке. Наконец, сеанс просмотра в лаборатории неизбежно краток, поэтому любые обобщения и выводы, касающиеся ситуации длительных, ежедневных просмотров, должны носить исключительно гипотетический характер.

Определенные ограничения налагаются и на зависимую переменную. Бить куклу Бобо — не то же самое, что бить другого ребенка, как и мнимое причинение боли невидимому ребенку, находящемуся в другой комнате. По крайней мере частично по этическим соображениям в качестве лабораторного критерия агрессии обычно используется псевдоагрессия — не столь ярко выраженная межличностная агрессия, как реальная и не столь часто вызывающая негативную реакцию у взрослых. Прогностичны ли такие действия в отношении истинно агрессивного поведения в контексте значительно больших ограничений — спорный вопрос. Даже когда проблема этичности исследования не стоит, в лабораторных условиях трудно воссоздать сложное социальное поведение. Как мы увидим вновь в следующих главах, лабораторные аналоги поведения, интересующего специалистов в области психологии развития, часто довольно далеки от своих прототипов из реальной жизни. В некоторых случаях сама суть изучаемого вопроса исключает возможность лабораторного исследования. Если нас, к примеру, интересует агрессия дошкольников конкретно на игровой площадке и направленная именно на сверстников, то исследование в лабораторных условиях просто не входит в число возможных вариантов.

Кроме того, в лабораторных исследованиях особенно часто встают проблемы реактивности и установки на определенный ответ, о которых говорилось в предыдущей главе. В обстановке лаборатории ребенок может оробеть и замкнуться, может пытаться делать то, чего, как ему кажется, хочет от него взрослый, его вниманием может полностью завладеть загадочное оборудование и т. д. Действительно, как говорилось в главе 5, экспериментальные мероприятия можно завуалировать и тем самым свести к минимуму реактивность. Стоит также отметить, что лаборатории, используемые в большинстве исследований школьников, — это незанятые школьные кабинеты, которые нельзя назвать «лабораториями» в буквальном смысле этого слова, необычными, напичканными аппаратурой помещениями. Тем не менее факт остается фактом: дети попадают в непривычную обстановку, незнакомый взрослый создает условия, в которых они никогда не оказывались, и все эти отступления от естественного хода событий могут привести к разного рода систематическим ошибкам.

Мы не будем рассматривать по порядку все клетки табл. 6.1, а обратимся к той из них, которая более всего отличается от только что описанного случая — к исследованию типа «полевые условия — полевые условия». Проанализировав классические лабораторные и полевые исследования, мы легко заполним оставшиеся клетки таблицы.

План 4: «полевые условия—полевые условия»

Теперь нас интересуют исследования, в которых и манипуляции с независимой переменной, и измерение зависимой переменной осуществляются в естественных условиях. Действительно, как отмечалось ранее, экспериментальные манипуляции и измерения неизбежно в той или иной мере нарушают «естественность» обстанов-

ки. Тем не менее в этом случае в отличие от лабораторного исследования отправным пунктом является естественная среда. И если рассматривать обстановку проведения исследования на континууме от лабораторных к естественным условиям, то она, скорее всего, окажется ближе к естественной.

Примером изучения связи насилия на телеэкране и агрессии может служить исследование Фешбах и Сингер (Feshbach & Singer, 1971). Их испытуемыми были мальчики предподросткового и подросткового возраста, проживавшие в школьных общежитиях. Фешбах и Сингер имели возможность в течение шести недель контролировать режим просмотра телепрограмм своими испытуемыми. Одну половину мальчиков произвольно включили в группу с 6-недельным режимом просмотра агрессивных телепрограмм, а другую половину — в группу просмотра нейтральных телепрограмм. Таким образом, здесь манипуляции с просмотром телепередач проводились в естественной для испытуемых среде. Об эффекте режимов просмотра судили по рейтинговым оценкам естественной агрессии мальчиков за 6-недельный период. Эти оценки давали учителя и супервизоры. То есть зависимая переменная также находилась в рамках естественных условий.

Главное преимущество полевого исследования сосредоточено в слове «естественность». Более всего при изучении связи между насилием на телеэкране и агрессией нас интересует вопрос, воздействует ли агрессия, которую дети видят дома на телеэкране, на уровень их агрессии дома, в школе или на игровой площадке — где бы естественным путем они ни оказались. Как мы узнали, лабораторное исследование может дать только косвенный ответ на этот вопрос, поскольку в нем манипуляции с режимом просмотра телепрограмм производятся не в домашней обстановке, а уровень агрессии измеряется не в естественных условиях. В полевом же исследовании присутствуют именно те ситуации и именно то поведение, которое, собственно, и интересует исследователя. Это означает, что здесь внешняя валидность выше, чем в аналогичном лабораторном исследовании. Если мы действительно сможем осуществить необходимый контроль независимой переменной и если мы действительно сможем точно измерить зависимую переменную, то внешняя валидность также будет высока.

Недостатки полевого исследования зашифрованы в двух «если» предыдущего предложения. Некоторые манипуляции и измерения в естественных условиях произвести трудно, если вообще возможно. Задумайтесь на мгновение о проблеме контролирования в течение длительного периода времени того, что смотрит по телевизору огромное количество детей. Теперь вас не удивит столь малое число полевых исследований, подобных исследованию Фешбах и Сингер, или то, что эти исследования проводятся в основном в закрытых учреждениях, например в школах-интернатах. Отражают ли результаты, полученные в таких условиях и на таких популяциях, картину более привычной домашней обстановки — спорный вопрос. Кроме того, экспериментальный контроль сам по себе вносит существенные изменения в среду. Редко у кого из детей просмотр телевизионных передач находится под абсолютным контролем взрослого, и внезапное наложение такого контроля связано с риском разнообразных эффектов реактивности и смещения. В исследовании Фешбах и Сингер, к примеру, были получены данные, свидетельствующие о том, что мальчики, которым показывали исключительно нейтральные

телепрограммы, испытывали фрустрацию в связи с потерей возможности смотреть любимые передачи и по этой причине стали более агрессивными.

Следует отметить, что уровень значимости рассмотренных выше проблем в определенной мере зависит от типа независимой переменной. С некоторыми переменными производить естественные, не вызывающие эффекта реактивности манипуляции в полевых условиях легче, чем с просмотром телепрограмм. Не представляет труда при изучении зрительного внимания младенцев подвесить игрушки над кроваткой малыша (например, Weizmann, Cohen, & Pratt, 1971), или при изучении нравственного развития проконтролировать факторы ситуации, подталкивающие к мошенничеству в классной комнате (например, Hartshorne & May, 1928). С другой стороны, некоторыми переменными манипулировать даже сложнее, чем просмотром телепрограмм. Иногда практические и этические соображения делают экспериментальные манипуляции невозможными; методы воспитания — характерный и весьма важный пример. В других случаях исследователя интересует действие довольно специфических и жестко контролируемых стимулов, которые искусственны по своей сути — к примеру, монотонные звуки при изучении слуховой габитуации, образы, появляющиеся на доли секунд на экране тахистоскопа при изучении зрительного обнаружения, список слов при изучении кратковременной памяти. Иногда есть возможность поместить эти стимулы в естественную среду, однако это не имеет смысла: естественная среда тут же превращается в «неестественную», а другие факторы ситуации могут серьезно исказить результаты и вызвать смешение различий между испытуемыми. В таких случаях разумнее всего выбрать лабораторное изучение.

Вторая, общая для всех полевых исследований проблема касается второго из двух «если», о которых говорилось выше: точного измерения зависимой переменной. Рассмотрим в качестве примера слуховую габитуацию. Термин *габитуация* означает угасание ориентировочной реакции на стимул при его повторном действии. О таком привыкании, как правило, судят по изменению частоты сердечных сокращений при неоднократном предъявлении стимула. Очевидно, что частота сердечных сокращений — это не та зависимая переменная, которую можно измерить в полевых условиях. Даже если исследователю удастся принести в дом испытуемого электрокардиограф, все эти провода, электроды и т. п. сделают обстановку далекой от естественной. Тот же аргумент справедлив во всех ситуациях, когда зависимой переменной является физиологическая реакция.

Справедливость этого аргумента не ограничивается ситуацией измерения физиологических показателей. Целый ряд поведенческих реакций, интересных для специалистов в области психологии развития, вызвать и измерить в естественных условиях весьма затруднительно. В эту категорию входят и многие феномены, которые изучал Пиаже. Вполне вероятно, что представления о сохранении являются важнейшим компонентом мышления ребенка, и именно такова точка зрения Пиаже. Однако в естественном поведении представления ребенка о сохранении редко бывают выражены непосредственно и отчетливо; чтобы судить об их наличии или отсутствии необходим специальный тест, например тот, который иллюстрирует рис. 5.1. Действительно, как будет сказано в главе 11, есть определенный смысл максимально приблизить условия этого теста к естественным. Однако какие-то

действия для выявления этих представлений все же необходимы и приближают исследование к лабораторному, если рассматривать континуум от лабораторных к полевым условиям.

Суть вышесказанного состоит в том, что некоторые реакции в буквальном (например, частота сердечных сокращений) или переносном (например, понимание принципа сохранения) смысле не подлежат наблюдению, и поэтому их нельзя измерить в естественных условиях. А что если рассмотреть агрессию? Агрессия, в конечном счете, часто встречающаяся, явная, наблюдаемая и, по существу, социальная поведенческая реакция и, будучи таковой, с полным правом претендует на измерение в полевых условиях. И действительно, измерение агрессии в естественных условиях, как мы увидим, имеет ряд конкретных преимуществ перед попытками измерить агрессию в лаборатории. Однако актуальными остаются важнейшие проблемы — осуществимости и точности. Можно выделить два основных подхода к измерению социального поведения в полевых условиях: рейтинговые оценки поведения ребенка, которые дают те, кто хорошо знает (как в исследовании Фешбах и Сингер), и непосредственное наблюдение поведения. О методах наблюдения речь шла в главе 4, а в главе 12 методы наблюдения и рейтинговой оценки рассматриваются в контексте обсуждения частных аспектов социального развития. Тогда же станут понятными и сложности, связанные с обоими методами измерения. Пока же ограничусь тем, что еще раз подчеркну главное: какого бы рода ни было поведение, его точное измерение в полевых условиях весьма затруднительно.

План 2: «лаборатория—полевые условия»

В этом случае независимой переменной манипулируют в лаборатории, а зависимую переменную измеряют в полевых условиях. Примером изучения связи между насилием на телеэкране и агрессией может служить исследование, которое провели Эллис и Секира (Ellis & Sekyra, 1972). Первоклассников разделили на три группы в соответствии с тремя видами экспериментальных ситуаций: просмотр агрессивного мультфильма в небольшом не используемом школьном кабинете; просмотр нейтрального мультфильма в том же кабинете; и отсутствие просмотра. Таким образом, независимой переменной манипулировали в лабораторных условиях. Затем детей наблюдали в их обычных классах и регистрировали все акты агрессии. То есть зависимую переменную измеряли в полевых условиях.

Достоинства таких исследований сочетают в себе достоинства манипуляций в лабораторных условиях и измерений в полевых условиях, о которых говорилось выше. Эти исследования позволяют совместить достаточный контроль независимой переменной с экологической валидностью измерения зависимой переменной. Таким образом, укрепляется уверенность в том, что полученный эффект можно отнести на счет экспериментальных манипуляций, а также вероятность того, что этот эффект имеет место в реальной жизни. Кроме того, сочетание лабораторных манипуляций с полевыми измерениями делает возможным пространственное и временное разделение независимой и зависимой переменных, что маловероятно в исследованиях типа «лаборатория—лаборатория» или «полевые условия — полевые условия». Это разделение само по себе расширяет возможности генерализации результатов, а также снижает вероятность реактивности и искажения ответов.

В упомянутом исследовании (Ellis & Sekyra, 1972), к примеру, факт измерения агрессии, не в том же месте, где происходил просмотр мультфильмов, и не в то же время, снижает риск того, что подражание было обусловлено непосредственными ситуационными сигналами или тем, что, по мнению ребенка, взрослый ожидал от него подражания.

О слабых сторонах плана «лаборатория—полевые условия» также говорилось при рассмотрении лабораторных и полевых исследований. При помещении независимой переменной в лабораторию появляется вероятность искусственности и невозможности генерализации; а помещение зависимой переменной в полевые условия делает проблематичным точное измерение. Кроме того, иногда совмещению в одном исследовании лабораторного и полевого компонентов препятствуют трудности практического характера. Как мы вскоре увидим, исследования типа «лаборатория — полевые условия» в действительности составляют лишь малую долю исследований в психологии развития.

План 3: «полевые условия—лаборатория»

Последний вариант — манипуляции с независимой переменной в полевых условиях при измерении зависимой переменной в лаборатории. Примером служит исследование Парке, Берковитц, Лайенс, Уэст и Себастиан (Parke, Berkowitz, Leyens, West, & Sebastian, 1977). Исследование Парке и его коллег имеет много общего с описанным ранее исследованием Фешбах и Сингер (Feshbach & Singer, 1971). Их испытуемыми были мальчики-подростки, половине из которых в их общезитиях в течение 5 дней показывали фильмы с эпизодами насилия, а другой половине показывали фильмы нейтрального содержания. Однако в этом случае одной из методик оценки эффекта просмотра был лабораторный тест, в котором испытуемому предоставлялась возможность воздействовать разрядами электрического тока на невидимого сверстника (прием, сходный с тем, который использовался в исследовании Либерт и Бэрона (Liebert & Baron, 1972)). Таким образом, для оценки эффекта манипуляций с просмотром телепрограмм в естественной среде использовались условия лаборатории.

Опять-таки, сильные и слабые стороны данного подхода следуют из общих замечаний, касающихся лабораторных и полевых исследований. Естественность условий, в которых находится независимая переменная, — это достоинство, равно как и точность, с которой можно измерить зависимую переменную. С другой стороны, попытка ввести экспериментальный контроль в естественную среду может обернуться некоторой потерей и естественности, и контроля. Результатом же измерения зависимой переменной в лабораторных условиях может быть точный, но искусственно полученный показатель, неопределенным образом связанный с агрессией в реальной жизни. И вновь при попытке соединить в одном исследовании лабораторные и полевые условия могут возникнуть практические трудности.

Анализ и оценка

В нашем обсуждении различных условий проведения исследования постоянно присутствуют две темы. Первая касается поиска компромисса между разнообразными целями, которые хотел бы достичь исследователь. Нам, конечно, хотелось бы

получать из исследования ясные выводы о причинно-следственных связях, выводы, которые можно было бы перенести на широкий диапазон ситуаций реальной жизни. Это предполагает, что мы жестко контролируем наши независимые переменные, точно измеряем наши зависимые переменные и имеем контекст исследования, достаточно близкий к естественному, чтобы позволить сделать широкие обобщения. Именно такого сочетания очень трудно добиться, поскольку методологические приемы, работающие на достижение одной цели, зачастую препятствуют достижению других целей.

Вторая тема является следствием первой. Поскольку нет такого подхода к исследованию некоей проблемы, который был бы идеален, использование лишь одного подхода не бывает оправданным. Значительно больше информации дает применение в одном или в нескольких исследованиях разнообразных методов изучения. Такой комплексный подход, как мы уже знаем, называется методом *конвергентных операций*. Его смысл заключается в том, что слабые стороны любого метода исследования могут в той или иной мере компенсироваться сильными сторонами другого и что выводы, основанные на сопоставлении данных, полученных разными способами, более достоверны, чем выводы, основанные на данных, полученных с помощью лишь одного метода.

Применим данный аргумент к проблеме связи насилия на телеэкране и агрессии. Мы увидели, что каждый из методов изучения этого вопроса, взятый в отдельности, имеет ряд недостатков. Однако обнаружив (и в действительности это обнаружено), что использование всех описанных на предшествующих страницах подходов привело к единым результатам, — а именно, что просмотр телепрограмм с эпизодами насилия ведет к повышению агрессии, — мы смогли бы быть более спокойными за их валидность.

До этого момента, описывая условия проведения исследования, мы говорили в основном о том, что в принципе возможно. Обратимся теперь к реальным фактам и данным о частоте использования рассмотренных выше подходов на практике. Но прежде необходимо указать на некоторые сложности в самой классификации подходов, поскольку многие исследования нельзя однозначно или исключительно отнести к какой-то из четырех категорий, указанных в табл. 6.1.

Начнем с независимой переменной. Мы рассмотрели те условия, в которых производятся манипуляции с независимыми переменными. Однако в некоторых случаях характер переменной исключает возможность манипуляций. К таким переменным относят *субъектные переменные*, о которых шла речь в главе 2, — возраст, пол, раса, контролируемые и изучаемые через отбор испытуемых, а не через экспериментальные манипуляции. В подобных исследованиях зависимая переменная может оцениваться и в полевых, и в лабораторных условиях, однако в отношении независимой переменной разграничение на лабораторные и полевые условия уже неприменимо.

Рассмотрим теперь сложности, связанные с зависимой переменной. Если выбранным критерием агрессии является количество агрессивных воздействий на куклу Бобо в некоей экспериментальной комнате, исследование однозначно попадает в категорию лабораторного. Если об агрессии судят по агрессивным воздействиям ребенка на других детей на игровой площадке детского сада, исследование

однозначно попадает в категорию полевого. Однако как расценивать ситуацию, когда об агрессии судят со слов человека, которому хорошо известно типичное поведение ребенка — то есть используют метод интервью или рейтинга? Рассматривая исследование Фешбах и Сингер (Feshbach & Singer, 1971), мы отнесли данный тип исследования к категории «полевых», обосновав это тем, что поведение оценивается в естественной среде, а не в структурированных условиях лаборатории. Однако у такого показателя, основанного на вербальных отчетах, отсутствует непосредственность и естественность, подразумеваемые под термином «полевое исследование». По-видимому, он находится где-то между полевыми и лабораторными исследованиями: ближе к полевым в том смысле, что анализируется действительно естественное поведение; и ближе к лабораторным в том смысле, что операция измерения очевидна и направлена не прямо на фактическое поведение.

Последняя сложность касается разграничения независимых и зависимых переменных. В некоторых исследованиях оно отсутствует, поскольку исследование не предполагает ни манипуляций с экспериментальной переменной, ни выбора некоторой субъектной переменной. Цель исследования может состоять в описании, попытке открыть интересные феномены, которые в дальнейшем могут изучаться экспериментально. Или же целью может быть оценка испытуемых по двум или более переменным и выявление отношений между этими показателями, то есть *корреляционное исследование*, которое рассматривалось в главе 3. В любом случае должно провериться измерение, в лабораторных, в полевых условиях или при использовании сочетания этих двух подходов.

В обзоре журнальных статей (см. табл. 1.3) учитывался тип условий, в которых проводилось исследование, и степень контроля переменных. В связи с описанными выше сложностями, а также с ограниченной выборкой исследований, из полученных цифр не следует делать далеко идущие выводы. Тем не менее некоторые результаты заслуживают внимания.

Из проанализированных исследований 34 % включали те или иные манипуляции с независимой переменной (интересно отметить, что это значительно меньше 70 % — показателя из обзора, представленного в первом издании этой книги). 60 % этих исследований включали ту или иную субъектную переменную; наиболее типичной переменной был возраст, за которым по частоте следовал пол. В 55 % проанализированных исследований субъектные переменные были единственными независимыми переменными. Корреляционные исследования в чистом виде — то есть исследования, в которых не осуществлялись манипуляции с экспериментальной переменной и не подбиралась определенная субъектная переменная, — соответствовали 11 % статей. Кроме того, 37 % исследований с независимыми переменными включали тот или иной корреляционный компонент.

Что можно сказать о разграничении между лабораторными и полевыми условиями? Рассмотрим сначала исследования, в которых присутствуют некоторые манипуляции с независимой переменной. В 97 % этих случаев экспериментальные манипуляции производились в лабораторных условиях, оставляя лишь 3 % на манипуляции в полевых условиях. Если говорить об измерении результирующих переменных, в 48 % проанализированных исследований они проводились в лабораторных условиях, в 5 % все измерения проводились в полевых условиях, в 21 % все

измерения происходили при помощи интервью или опросников, а в 25 % использовалось сочетание разных подходов. Наконец, если обратиться к таблице 6.1, можно отметить, что план 1, «лаборатория—лаборатория», использовался наиболее часто — в 97 % исследований из тех, где независимая переменная подвергалась экспериментальным манипуляциям. План 4, «полевые условия — полевые условия», использовался в оставшихся 3 %.

О чем говорят эти цифры? Данный обзор подтверждает уже сложившееся мнение о том, что в исследованиях в области психологии развития имеется тенденция к использованию манипулятивных планов и лабораторных условий в ущерб натуралистическому и эксплоративному подходам. Такое особое внимание к лабораторным методам часто становится поводом для критики. Макколл (McCall, 1977), например, пишет, что «мы редко удерживаемся от экспериментального изучения настолько, чтобы, наблюдая поведение в естественных условиях, успеть сделать его систематическое описание» (р. 336). Бронфенбреннер (Bronfenbrenner, 1977) утверждает, что «большая часть современной психологии развития — это наука о *нетипичном поведении ребенка в нетипичных ситуациях с незнакомыми взрослыми в течение чрезвычайно коротких промежутков времени*» (р. 513, курсив оригинала).

То, что лабораторные исследования имеют свои ограничения, понятно из того, что уже было сказано в этой главе. Однако смысл приведенной выше критики в том, чтобы поставить под сомнение не ценность контролируемого лабораторного исследования, а непропорциональное внимание к этой стратегии в ущерб другим способам изучения детей. Суть опять-таки в необходимости использования конвергентных операций. Поскольку любой метод исследования в отдельности несовершенен, любую научную проблематику нужно изучать как можно более разнообразными методами.

В концепции, к которой мы сейчас обратимся, также рассматриваются объекты особого внимания психологического исследования, однако с несколько иных позиций. Эта концепция послужит связующим звеном между темой условий проведения исследования и темой целей, так как включает проблематику обеих этих тем. Речь идет о теории *экологических систем* Ури Бронфенбреннера (Urie Bronfenbrenner, 1977, 1979, 1989, 1993).

Экологические системы

В концепции экологических систем основное внимание уделяется контексту, в котором проходит развитие, и взаимосвязям между разными контекстами, играющими роль в развитии ребенка. В предыдущем разделе тоже, конечно, говорилось о контексте, однако в несколько ином смысле. При обсуждении типов условий нас интересовал вопрос, где проходит *исследование*, если в общем виде разделить условия на лабораторные и полевые. Однако, по определению, *развитие* протекает всегда в полевых условиях, в различных типах естественной среды (дома, в школе, на игровой площадке и т. д.), в которой проходит жизнь ребенка. Вопрос, которому посвящены работы Бронфенбреннера, заключается в том, каким образом концептуализировать и изучать эти разнообразные контексты или «экологические системы».

Начнем изложение рассуждений Бронфенбреннера с рассмотрения контекстов, упомянутых в предыдущем абзаце — дома, школы, игровой площадки. Каждый из них, разумеется, включает не только некие физические параметры, но и характерные виды деятельности, а также важнейших социальных агентов — родителей, братьев и сестер, учителей, сверстников. Такие контексты Бронфенбреннер называет **микросистемой**. Микросистема — ближайшая из экологических систем, слой, с которым ребенок находится в непосредственном контакте, условия, прямо воздействующие на ребенка. В категорию микросистемы попадает взаимодействие с членами семьи в домашней обстановке. Сюда же относятся общение с друзьями на игровой площадке, беседы с учителями в школе, проповедь, прослушиваемая в церкви.

Как мы вскоре отметим, большинство исследований в психологии развития направлено на изучение именно микросистемы; поэтому не составляет труда подобрать соответствующие примеры. В действительности, большинство приведенных в этой книге примеров демонстрирует изучение микросистемы. Это справедливо и в отношении проблемы, рассмотренной выше для пояснения вопроса условий проведения исследования: проблемы связи между просмотром агрессивных телепрограмм и агрессивным поведением. При изучении возможного влияния просмотра телепрограмм с эпизодами насилия на уровень агрессии нас интересует эффект непосредственного личного опыта: как то, что ребенок видит и слышит, сидя перед телевизором, отражается в дальнейшем на его поведении? Если рассматривать эту проблему в более общем виде, при исследовании влияния моделей поведения — телевизионных или любых других — на определенный аспект развития ребенка нас интересует именно микросистема: воздействие личного опыта, получаемого при участии социальных агентов из мира ребенка. Для изучения микросистемы необходимо либо включить в исследование реальных социальных агентов (родителей, братьев, сестер и т. д.), либо создать лабораторный аналог, в котором экспериментатор будет выступать в роли соответствующих фигур из реальной жизни.

Интерес психологов к микросистеме, безусловно, понятен — это тот слой среды, в котором действуют каузальные факторы и происходит развитие ребенка. Вопрос в принципе в том, почему нас должно интересовать что-то еще, то есть какие другие слои или системы должны дополнить микросистему для создания полной картины контекста развития? Рисунок 6.1 в символической форме иллюстрирует ответ Бронфенбреннера. Как явствует из рисунка, теория экологических систем подразумевает наличие помимо микросистемы еще трех контекстуальных слоев. **Мезосистема** — это система отношений между микросистемами. Она может включать, к примеру, участие родителей в процессе школьного обучения ребенка или взаимодействие его братьев и сестер с его друзьями. Под термином **экзосистема** подразумеваются социальные системы, которые могут влиять на ребенка, но в которых он не принимает непосредственного участия. Попечительский совет, утверждающий школьные правила, которые имеют отношение к ребенку, — элемент экзосистемы; к ней же относится работодатель, предоставляющий (или не предоставляющий) матери ребенка декретный отпуск. Наконец, **макросистема** — это культура или субкультура, в которой развивается ребенок. Обычно макросистема

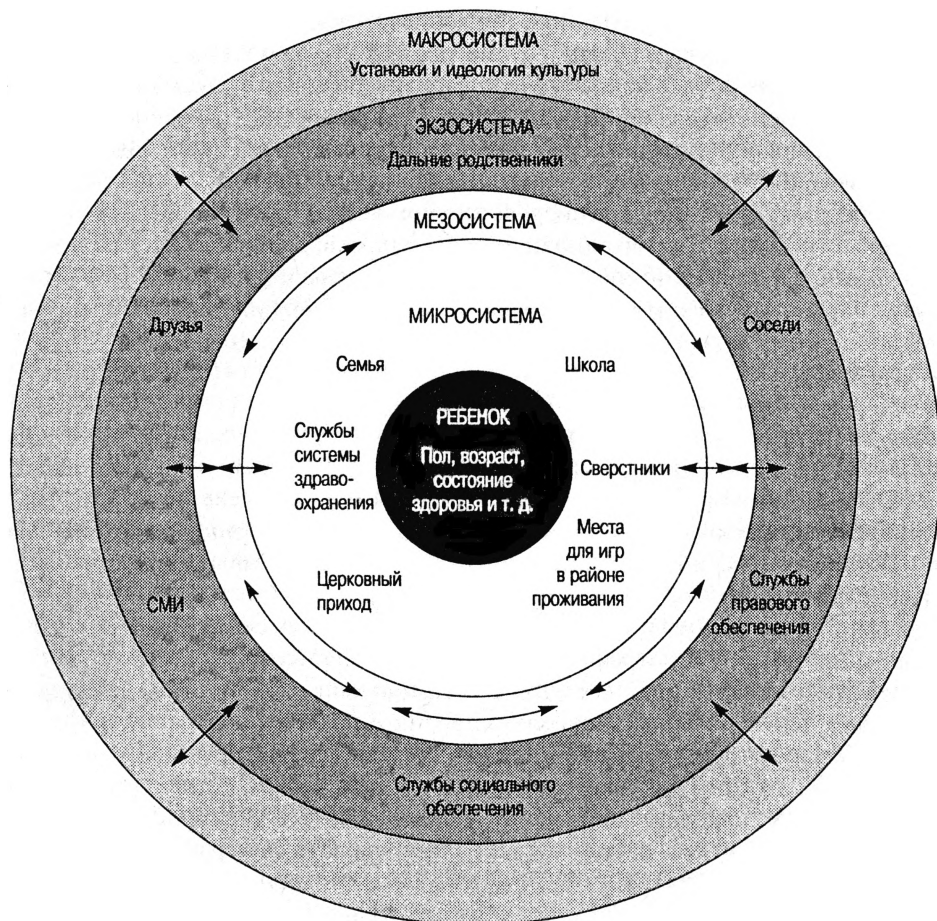


Рис. 6.1. Модель экологических систем Бронфенбреннера. («Sociocultural Risk: Dangers of Competence» J. Garbarino, 1982 In C. B. Kopp & J. B. Krakow (Eds.) *The Child: Development in a Social Context* (Fig.12-1 from p. 648) Reading, MA: Addison-Wesley)

включает в себя подсистемы. К примеру, для ребенка макросистема может состоять не только из культуры США в целом, но и из специфической культуры крупного северо-восточного города и, вполне вероятно, из этнической общности, населяющей определенный район этого города.

Что следует из концепции экологических систем? В частности, каковы *методологические* следствия — какие способы проведения и интерпретации исследования предлагает эта теория? По мнению Бронфенбреннера (Bronfenbrenner, 1979, p. 38), основной вывод из нее таков: «В экологическом исследовании главными эффектами скорее всего будут взаимодействия». Под этим Бронфенбреннер подразумевает, что влияние одной экологической системы может тем или иным способом корректироваться некой второй экологической системой. Это представление

об экологических системах как о системах, взаимосвязанных и взаимодействующих иллюстрируют концентрические круги на рис. 6.1. Мы можем (и обычно так и происходит) изучать только одну экологическую систему — чаще всего, микросистему. Однако системы *не существуют* по отдельности. Микросистема всегда включена в мезосистему, а микро- и мезосистемы уходят корнями в более высокие, глобальные слои.

Какого рода взаимодействия можно обнаружить при одновременном рассмотрении двух и более экологических систем? Здесь мы проанализируем лишь два примера; другие примеры, а также более подробное описание теории Бронфенбреннера и ее практического значения для исследований можно найти в трудах самого Бронфенбреннера и в работах, посвященных его теории (Moen, Elder, & Luscher, 1955; Pence, 1988).

Первый пример предваряет материал, более подробно изложенный в главе 12. Одна из тем, освещенных в этой главе, — изучение используемых родителями методов воспитания, а один из описанных там подходов — исследовательская программа Баумринд (Baumrind, 1971, 1989). Баумринд выделяет четыре стиля воспитания, один из которых мы рассмотрим здесь. *Авторитарным* Баумринд называет способ социализации, для которого характерно частое использование угроз и наказаний, властность родителей, относительно редкое использование ими объяснений и недостаток родительской теплоты. Кратко описанный данным образом, авторитарный стиль представляется не слишком удачным подходом к воспитанию детей, и действительно, так и есть. По данным Баумринд, авторитарный стиль стабильно связан с относительным неблагополучием детей. Эти результаты были воспроизведены и другими исследователями, использовавшими разные методики и изучавшие разные возрастные группы (например, Lamborn, Mounts, Steinberg, & Dornbusch, 1991).

Однако все упомянутые выше исследования проводились преимущественно на выборках белых представителей среднего класса. Следовательно, они ограничивались одной макросистемой. Как оказалось, это ограничение имеет значение. Дорнбуш и коллеги (Dornbusch, Ritter, Leiderman, Roberts, & Fraleigh, 1987) проанализировали влияние описанных Баумринд стилей на выборках подростков и их родителей из разных этнических групп (белых, афро-американцев, азиатов и латиноамериканцев). Хотя в целом результаты были близкими к результатам, полученным в других подобных исследованиях, прогнозировавшиеся связи между методами воспитания и уровнем благополучия ребенка оказались максимально выраженными в белых семьях; во всех других этнических группах наблюдались некоторые исключения и специфические моменты. Особенно интересны в этом плане азиатские семьи. Для них характерен высокий уровень авторитарности, однако у азиатских подростков оказалась высокой и школьная успеваемость (главный из оценивавшихся показателей благополучия ребенка). Таким образом, в рамках макросистемы азиатско-американской культуры авторитарный стиль воспитания по всей видимости несет более позитивный смысл, чем в изучавшейся до этого евро-американской культуре. Говоря о данном факте, Бронфенбреннер (Bronfenbrenner, 1993, p. 39) приводит слова одного азиатско-американского студента, в которых, по крайней мере частично, отражена суть указанного культурного разли-

чия. Отвечая на вопрос о том, что думают азиаты об авторитарном стиле воспитания, он сказал: «Именно так мы узнаем, что родители нас любят».

Пример со стилями воспитания иллюстрирует взаимодействие микро- и макросистем. Следующий пример демонстрирует значение экзосистемы. Элдер и коллеги (Elder, 1974; Elder & Caspi, 1990) изучали последствия Великой депрессии 30-х годов на развитие детей. С этой целью они использовали архивные данные, то есть лонгитюдные данные, собранные в 1930-х годах, которые были пересчитаны и заново проанализированы. Мы остановимся на том аспекте их работы, который касается четырех групп мальчиков: половина из них в начале Депрессии находилась в дошкольном возрасте, а другая половина — в периоде позднего детства; у половины мальчиков из каждой возрастной группы отец стал безработным, а у другой половины — не потерял работы. Неудивительно, что негативные эффекты были наиболее выражены в семьях, где отец потерял работу и которые поэтому испытывали большие экономические трудности. Однако более серьезно это отразилось на мальчиках, которые к началу Депрессии еще не достигли школьного возраста. В их семьях снижение дохода и изменение ролевых отношений привели к разнообразным изменениям во внутрисемейной жизни (например, непоследовательности дисциплинарных воздействий, повышению уровня напряженности, снижению авторитета отца как ролевой модели), что негативно сказалось на детском развитии. Мальчики, которые к началу Депрессии были старше, имели возможности принять на себя более взрослые роли и внесемейные обязанности и таким образом в некоторой степени избежали влияния изменений в семейной динамике. Интересно также отметить, что на девочках такая ситуация отражалась не столь негативно, и это не зависело от их возраста к началу Депрессии. Таким образом, исследование демонстрирует не просто влияние экзосистемы (занятости отца) на микросистему (семейные отношения), а взаимодействие этих двух систем. Ухудшение материального благополучия по-разному сказывалось на разных детях и на семьях разных типов.

Значение одновременно и возраста и пола ребенка в указанном исследовании иллюстрирует еще одно положение теории Бронфенбреннера, на которое следует обратить внимание. При знакомстве с изложенным выше материалом, с его акцентом на внешних силах, действующих на ребенка, могло создаться впечатление, что сам ребенок пассивен и его развитие зависит от того, в какую среду он попадет. Однако отнюдь не в этом суть экологического подхода. В действительности, Бронфенбреннер был одним из первых специалистов в области психологии развития, который стал защищать принцип, получивший сегодня общее признание: ребенок играет активную роль в своем собственном развитии, в определенной мере создавая среду, в которой он живет, а не просто реагируя на нее. Бронфенбреннер пишет (Bronfenbrenner, 1989), что по иронии судьбы и против его желания работы по экологическим системам возбудили чрезмерный исследовательский интерес к контексту без учета особенностей самого ребенка — то есть «контекст без развития» сменил предыдущее и равно ошибочное «развитие без контекста». Очевидно, что исследование должно охватывать и процесс развития, и контекст, в котором оно протекает.

Таблица 6.2

Формы прикладного исследования в психологии развития

Категория	Цель	Примеры
Оценка	Выявить значимые характеристики в группах риска нарушений в развитии	Оценочная шкала поведения новорожденного для оценки статуса новорожденных, переживших пренатальный стресс Использование шкалы интеллекта Векслера для дошкольников и младших школьников (WPPSI) или шкалы Стэнфорд—Бине для выявления возможного отставания в когнитивном развитии у дошкольников из неблагополучных семей
Вмешательство	Изменить среду в направлении создания условий, предупреждающих, корректирующих или уменьшающих нарушения в развитии	Проект <i>Head Start</i> и аналогичные программы, направленные на повышение уровня подготовленности к школьному обучению, когнитивной и социальной компетентности дошкольников, воспитывающихся в неблагоприятных условиях Программы тренировки памяти для лиц, испытывающих затруднения с памятью
Исследование социально значимой проблемы	Предоставить информацию, необходимую для решения значимых с практической точки зрения вопросов	Исследования, направленные на оценку точности свидетельских показаний ребенка в случаях подозрений на жестокое обращение Исследования, направленные на выявление детерминант употребления психоактивных веществ в подростковом возрасте
Вклад в государственную социальную политику	Проведение исследований с целью обеспечения научной обоснованности решений, принимаемых на государственном уровне	Использование результатов исследований в области психологии развития для выработки стандартов качества дошкольных учреждений Учет результатов исследований в области психологии развития при принятии решений об опеке над детьми и их посещении в случае развода
Популяризация психологических знаний	Сделать результаты исследований доступными для тех, кому они могут оказаться полезными	Издание книг и брошюр с практическими советами для родителей детей подросткового возраста Консультации психологов (например, в попечительских школьных комитетах, в суде, при организации деятельности приютов), когда необходимы психологические знания

Есть ли у нас такие исследования? При проведении контент-анализа журналов *Child Development* и *Developmental Psychology* исследования классифицировались по параметру экологической системы или экологических систем, которые они включали. Цифры, полученные в этом обзоре, следует интерпретировать с осторожностью по двум причинам: эта классификация неоднозначна, и разные аналитики

могли в некоторых случаях разойтись во мнении; кроме того, поскольку исследования, выходящие за пределы микросистемы, носят глобальный и комплексный характер, их результаты чаще всего публикуют в виде монографий или книг, а не статей в журналах. Помня об этих ограничениях, отметим, что в 37% исследований, к которым применима данная классификация, изучались только микросистемные процессы (чаще всего, влияние родителей на развитие детей). В большинстве оставшихся исследований также рассматривались аспекты микросистемы, однако уже в своем взаимодействии с другой или другими экологическими системами. Чаще всего изучалась мезосистема или взаимосвязи между микросистемами (например, влияние родительской социализации на отношения со сверстниками); анализ мезосистемы осуществлялся в 21% исследований. Оценка эффектов экосистемы проводилась в 10% исследований, а макросистемным процессам уделялось внимание в 14% исследований. Таким образом, эти цифры говорят о том, что исследователи предпринимают определенные попытки при изучении развития рассматривать разные экологические системы, даже в том объеме, в котором это можно отразить в журнальной статье. Однако микросистема остается наиболее типичным объектом изучения.

Прикладные исследования

Перейдем теперь от условий, в которых проводятся исследования, к их целям. Как отмечалось в самом начале этой главы, по своим целям исследования делятся на две категории: фундаментально-научные исследования и прикладные. Поскольку большая часть книги посвящена первой категории, задача этого раздела — рассмотреть некоторые вопросы, касающиеся второй категории исследований. Но прежде еще раз подчеркну то, о чем уже говорилось: фундаментальные и прикладные исследования в своей основе не имеют различий и основаны на единых методологических принципах. Поэтому в этом разделе не излагаются какие-то новые принципы сверх тех, которые уже были рассмотрены. Однако в нем будет рассказано о том, как эти принципы можно применить для решения важнейшей и сложной задачи проведения исследования, работающего на благо общества.

Что подразумевается под термином «прикладное исследование»? Фишер и Лернер (Fisher & Lerner, 1994b, p. 4) определяют прикладную психологию развития как «отрасль психологии, занимающуюся созданием условий, способствующих развитию и предотвращающих отставание в развитии». Далее они определяют несколько конкретных целей и связанные с ними формы исследования, входящие в сферу прикладной психологии развития. Эти цели и примеры типов исследований, связанные с каждой из целей, приведены в табл. 6.2. Таблица составлена частично по материалам Фишер и Лернер, а частично — по другим работам, посвященным вопросам прикладных исследований (например, Zigler & Finn-Stevenson, 1992). Все эти цели имеют огромное значение; однако по причине ограниченного объема книги мы остановимся лишь на двух формах прикладного исследования: оценке лиц, входящих в группу риска нарушений в развитии; и на программах вмешательства, направленных на предупреждение и коррекцию нарушений в развитии. Бо-

Таблица 6.3

Примеры тестов из ОШПН

Тест	Описание
Угасание реакции на свет	Направьте свет на глаза спящего ребенка, отметьте реакцию; после исчезновения реакции подождите 5 секунд и вновь направьте свет на глаза ребенка; проделайте это 10 раз либо до достижения габитуации
Угасание реакции на погремушку	Потрясите погремушкой возле уха спящего ребенка, отметьте реакцию; проделайте это 10 раз либо до достижения габитуации
Зрительная ориентировка на неодушевленный объект	Медленно передвигайте красный шарик в поле зрения ребенка; отметьте его способность следить за предметом в горизонтальном и вертикальном направлениях
Зрительная ориентировка на одушевленный объект	Медленно измените положение своего лица в поле зрения ребенка; отметьте его способность следить за лицом в горизонтальном и вертикальном направлениях
Слуховая ориентировка на одушевленный объект	Находясь вне поля зрения ребенка, тихо скажите ребенку что-нибудь на ухо; обратите внимание, определяет ли ребенок, с какой стороны звучит голос
Оборонительные движения	Накиньте пеленку на лицо лежащего ребенка, подождите 30 секунд; отметьте защитные движения (попытки сдвинуть)

Источник: *Neonatal Behavioral Assessment Scale* (3rd ed.) T. Berry Brazelton & J. K. Nugent, 1995; London: Mac Keith Press.

Более подробное описание этих и других видов прикладных исследований можно найти в специальной литературе (Fisher & Lerner, 1994a; Zigler & Finn-Stevenson, 1992). Обратите также внимание, что выпускаются два журнала, специально посвященных прикладной психологии развития: *Journal of Applied Developmental Psychology* и *Applied Developmental Science*.

Оценка

Проблема оценки может возникнуть на любом этапе жизненного пути — что, в действительности, справедливо по отношению ко всем формам прикладного исследования. Возьмем в качестве примера самую раннюю фазу жизни: период новорожденности. Особый интерес представляет возможность на самых ранних этапах развития выявить факторы риска нарушений и последствия их действия.

Какие обстоятельства связаны с риском для новорожденных? К сожалению, их достаточно много. Фактором риска является недоношенность, недостаток питания в пренатальный период. Угрозу может представлять и болезнь матери во время беременности, равно как и использование обезболивающих препаратов во время родов, а также употребление матерью в период беременности алкоголя и наркотиков.

Разумеется, наш интерес к возможным эффектам подобных факторов не ограничен периодом новорожденности; фактически, наибольшее беспокойство вызы-

вает возможность долгосрочных или перманентных нарушений. Кроме того, существует вероятность того, что некоторые последствия проявятся не сразу при рождении или вскоре после него, а гораздо позже. Тем не менее первым встает вопрос статуса малыша при рождении, и для ответа на этот вопрос нам нужен инструмент оценки, с помощью которого можно узнать, каково состояние новорожденного. До сих пор чаще всего с этой целью используют **оценочную шкалу поведения новорожденного (ОШПН)** (*Neonatal Behavioral Assessment Scale — NBAS*), разработанную более 30 лет назад и сегодня представленную в своей третьей версии (Brazelton & Nugent, 1995).

ОШПН — не первая попытка оценить статус новорожденного. Однако в тестах, созданных ранее, основной упор делался на оценку рефлекторной деятельности как показателя неврологической сохранности малыша. Так, отсутствие определенных рефлексов или сохранение определенных рефлексов на этапе развития, когда они должны исчезнуть, может быть признаком повреждения мозга. В ОШПН также основное внимание уделяется рефлекторной деятельности; при обследовании осуществляется проверка 20 рефлексов. Однако к этим рефлексам добавлен поведенческий компонент — 28 поведенческих тестов, выявляющих индивидуальные различия новорожденных в уровне адаптации. В табл. 6.3 представлено несколько поведенческих тестов. Как можно увидеть, шкала включает разнообразные показатели реакции *габитуации*, кратко описанной ранее в этой главе и более подробно рассмотренной в главе 10. Она также включает несколько показателей реагирования ребенка на социальные стимулы, которые обеспечивает взрослый, например, ориентировку на лицо или голос и возможность взрослого успокоить малыша. В действительности, внимание не только к самому ребенку, но и к его взаимодействию с другими является отличительной особенностью ОШПН. Еще одна особенность — выявление *оптимального* уровня реагирования ребенка — достаточно гибкое поведение обследующего, позволяющее раскрыть лучшее, на что способен новорожденный.

ОШПН обладает множеством качеств, свойственных всем оценочным методикам. Она включает широкий диапазон заданий; поскольку оценивается общий уровень компетентности новорожденного, важно, чтобы оценка не ограничивалась одним или несколькими типами поведения. Делается также акцент на *стандартизации* — создании условий, при которых процедура обследования и подсчета результатов была бы одинаковой для всех младенцев. Основание для стандартизации очевидно. Мы сможем проинтерпретировать результаты новорожденного, а также сравнить их с результатами других малышей, только будучи уверенными в том, что все новорожденные прошли одинаковую процедуру обследования. С другой стороны, как было указано выше, делается акцент и на *гибкость* — на проведение по мере необходимости адаптации четкой процедуры для выявления оптимального уровня реагирования малыша. Большинство оценочных методик дают возможность адаптировать процедуру, не выходя за рамки стандарта, по усмотрению обследующего; однако их различает то, насколько широки эти рамки. ОШПН, несомненно, можно назвать гибкой методикой. Необходимость гибкого подхода, в свою очередь, является следствием важнейшего факта — центральная фигура при проведении обследования это сам малыш; а малыши, о чем мы поговорим отдель-

но в главе 10, не всегда бывают добросовестными испытуемыми. При использовании ОШПН сложность процесса тестирования увеличивается, учитывая возраст недостаточно высокий уровень развития изучаемой группы, сложность может увеличиться даже более, если обследуемый новорожденный имеет некоторые отклонения в развитии. А следствием этого является то, что проводить обследование с использованием ОШПН может только квалифицированный специалист. Специальная подготовка и умение проводить предварительную клиническую диагностику требуется для использования любых оценочных методик; но в случае с ОШПН эти требования особенно высоки.

Дает ли ОШПН валидную оценку статуса новорожденного? В главе 4 рассматривались общие вопросы валидности теста. Валидизация ОШПН частично основана на каждом из описанных там параметров; однако первичной всё-таки является *критериальная валидность* — возможность с помощью данной методики спрогнозировать некие внешние показатели статуса развития. Критериальная валидность, в свою очередь, делится на текущую валидность (связь с фактическим статусом) и прогностическую валидность (связь с будущим статусом). Признаком текущей валидности является способность теста выявлять отклонения, уже имеющиеся в пренатальный период или при рождении у детей, подвергшихся воздействию неблагоприятных условий. Таким образом, если вернуться к вопросу, с которого я начал этот раздел, с помощью ОШПН можно удостовериться в том, что каждое из указанных ранее обстоятельств (недоношенность, употребление матерью психоактивных веществ и т. д.) действительно ведет к повышению риска нарушений в развитии. (Заметьте, однако, что в таких случаях нарушения не являются чем-то неизбежным; иначе говоря, нарушения могут возникнуть — и быть выявлены при помощи ОШПН — и при отсутствии известных факторов риска). Для установления прогностической валидности требуется проведение лонгитюдного исследования, в котором должно выясниться, можно ли по результатам ОШПН судить о дальнейшем развитии ребенка. Данные ряда исследовательских работ свидетельствуют о том, что статус при рождении прогностичен в отношении многих аспектов развития, в основном в период младенчества, однако в некоторой степени и в более поздние периоды (Brazelton, Nugent, & Lester, 1987). Конечно, прогностические возможности в любом случае далеки от идеальных, но это и неудивительно: никто не ждет, что на основе результатов обследования в первые дни жизни можно сделать однозначный прогноз дальнейшего развития. Удивительно, что эта методика вообще прогностична.

Из сказанного о валидизации ОШПН следует еще несколько замечаний. При использовании любой оценочной методики самое малое из того, что нас интересует, это наличествующий статус обследуемого. Когда объектом оценки является взрослый, мы можем удовлетвориться информацией о текущем состоянии. Однако при обследовании детей цель состоит не только в оценке имеющегося состояния, но и в прогнозе состояния на *будущее*. Как, к примеру, определенные нарушения в период новорожденности связаны с социальным развитием в годовалом возрасте? Неудивительно, что чем отдаленнее прогноз, тем меньше его точность. Кроме того, даже если мы не будем делать прогнозов, ни одна методика, по крайней мере психологическая, не дает полной и безошибочной картины изучаемого

конструкта. Все оценочные методики в той или иной мере несовершенны; ставя диагнозы и делая прогнозы, влияющие на жизнь людей, всегда нужно помнить об этих ограничениях. Наконец, работа с ОШПН демонстрирует пользу, которую при правильном использовании могут приносить оценочные методики. Во-первых, это польза для обследуемого. Оценка дает информацию об индивиде — зачастую, но, конечно, не всегда, касающуюся нарушений в развитии, — которая может помочь так организовать окружающую его среду, чтобы уменьшить негативные проявления и оптимизировать процесс развития. Во-вторых, это польза для изучаемой популяции в целом. Если, к примеру, при помощи методики, подобной ОШПН, удастся обнаружить, что некоторые действия родителей представляют угрозу для развития ребенка, тогда, с целью снижения риска для следующих поколений, может осуществляться просветительская работа, а также могут вводиться изменения в государственную социальную политику. (Более подробно рассмотренные здесь вопросы обсуждаются у Gibbs & Teti, 1990; Корр, 1994.)

Вмешательство

Оценка и вмешательство тесно связаны друг с другом. Как только что отмечалось, следующим за оценкой этапом является вмешательство, направленное на коррекцию обнаруженных нарушений. Иначе говоря, оценка — это компонент любого вмешательства: по меньшей мере нам нужны первичные показатели для определения исходного статуса обследуемых и конечные показатели для определения эффективности вмешательства.

За примером программы вмешательства обратимся к жизненному этапу, противоположному тому, о котором говорилось выше. Одним из характерных стереотипов, касающихся пожилого возраста, — и одной из распространенных жалоб пожилых людей — являются проблемы с памятью. Как мы узнаем из главы 13, это правило имеет ряд ограничений и исключений; тем не менее для большинства людей преклонного возраста трудности с памятью — это действительно чрезвычайно неприятная реальность. Вопрос, к которому мы сейчас обратимся, заключается в следующем: может ли вмешательство в форме программ тренировки памяти разрешить некоторые из этих трудностей?

В целом, как свидетельствует ряд исследовательских работ по проблеме тренировки памяти (Verhagen, Marcoen, & Goossens, 1992; West, 1995), ответ утвердительный. Сначала я приведу наглядный пример программы вмешательства, а затем мы рассмотрим некоторые общие вопросы, имеющие отношение к таким программам. Шейх, Хилл и Йесэвидж (Sheikh, Hill, & Yesawage, 1986), оценивали возможность тренировки памяти в группе пожилых людей, которые либо сами вызвались принять участие в исследовании в целях улучшения памяти, либо были направлены врачом или домом для престарелых. Вид изучавшейся ими памяти имеет огромное значение в любом возрасте — способность фиксировать в памяти связь между именем и лицом. Первичная оценка показала, что у многих испытуемых действительно наблюдаются трудности в такого рода запоминании: средние показатели протеста составляли лишь 4,13 правильных пар «имя—лицо» из 12 возможных. Последующий курс тренировки памяти, проводившийся в группе и длившийся 14 часов, был посвящен обучению *мнемоническим стратегиям* — то есть

методам облегчения запоминания и воспроизведения. Как будет указано в главах 11 и 13, результаты ряда исследований свидетельствуют о том, что стратегии запоминания во многом определяют индивидуальные и возрастные различия в уровне памяти; неудивительно поэтому, что эти стратегии являются главным компонентом большинства программ тренировки памяти. В исследовании Шейх с соавторами (Sheikh et al., 1986) (как, фактически, во многих подобных программах) особый акцент делался на использовании воображения — к примеру, запоминание фамилии Уэйлен (Whealen) посредством представления кита (*whale*) в огромном рту мистера Уэйлена. Обучение использованию подобных приемов оказалось эффективным: во всех экспериментальных группах наблюдалось почти 100% улучшение по сравнению с результатами протеста, а также более высокие результаты, чем у контрольной группы, не проходившей обучения. Кроме того, улучшение после прохождения наиболее эффективных форм тренировки наблюдалось и при повторном тестировании 6 месяцев спустя.

Исследование Шейх с соавторами (Sheikh et al., 1986) иллюстрирует ряд вопросов, имеющих отношение ко всем программам тренировки памяти, а также к программам вмешательства в целом. Один из вопросов касается *выявления испытуемых* — на кого будет направлено экспериментальное воздействие? По этому параметру отличаются и программы улучшения памяти для пожилых людей: в одних случаях воздействие направлено на пожилых людей в целом, отобранных лишь по критерию возраста и желания принять участие в исследовании; в других случаях в группу набирают только тех пожилых людей, у которых действительно наблюдаются нарушения памяти; в третьих случаях группа ориентирована клинически, и здесь реализуется попытка смягчить некоторые проблемы с памятью, являющиеся проявлением патологических изменений (например, при болезни Альцгеймера). В других сферах вмешательства программы, как правило, направлены на определенные подгруппы популяции — к примеру, не на дошкольников вообще, а на тех, у кого имеются расстройства внимания, или не на подростков вообще, а только на тех, кто когда-либо употреблял наркотики. Именно в таких случаях особое значение приобретает оценка: нам нужна методика, с помощью которой мы сможем выявить лиц, нуждающихся в помощи, а также определить конкретную область нарушений с тем, чтобы составить соответствующую программу вмешательства.

Вслед за выявлением потенциальных испытуемых встает вопрос экспериментального плана. Схемой большинства проектов вмешательства является последовательность, которую в главе 3 мы обозначили как «Одна группа претест/посттест»: исходная оценка, за которой следует вмешательство, а затем — конечная оценка. Как говорилось в главе 3, существенное различие между результатами пре- и посттестов не всегда свидетельствует о пользе программы; при использовании этого плана часто возникают альтернативные объяснения или искажения валидности выводов о причинно-следственных связях. По этой причине, осуществляя такие проекты, многие исследователи (как и Шейх с коллегами) набирают контрольную группу лиц, не подвергающихся экспериментальному воздействию. Контрольная группа дает возможность оценить вклад посторонних по отношению к экспериментальному воздействию факторов (например, эффект практики или ре-

грессии к среднему) и, таким образом, более точно определить, какая доля улучшения обусловлена вмешательством. Однако, как это часто бывает, когда дело касается исследовательского плана, проведение методологически грамотной процедуры осложняется наличием некоторых проблем практического характера. Испытуемые, к примеру, могут возражать против участия в контрольной группе, тем самым создавая вероятность систематической ошибки при отборе. Есть также и некоторые этические моменты, связанные с отказом нуждающимся людям в потенциально благотворном воздействии ради экспериментального контроля; этот вопрос обсуждается в главе 8 под рубрикой *отказ в экспериментальном воздействии*. В связи с этими трудностями во многих проектах вмешательства по завершении формального сбора данных контрольной группе предлагается того или иного рода помощь. В исследовании Шейх с соавторами (Shekh et al., 1986), к примеру, по окончании исследования контрольная группа прошла 4-часовой курс обучения мнемоническим стратегиям.

Как отмечалось, точная оценка важна не только для выявления испытуемых по результатам протеста, но и для определения эффекта вмешательства по результатам посттеста. Очевидно, что нас интересует улучшение, следующее за вмешательством — улучшение, выходящее за рамки того, что можно было бы объяснить действием других факторов (отсюда необходимость контрольной группы), и улучшение, имеющее не просто статистическую, а практическую значимость. Что конкретно имеет смысл ожидать, зависит от характера проекта. Однако при оценке эффективности вмешательства, в том числе и программ тренировки памяти, следует иметь в виду два общих соображения. Первое касается *переноса*, или *генерализации*: распространяется ли польза воздействия не только на конкретные задания, стимульный материал и контекст, использовавшийся в программе? В исследовании Шейх с соавторами, например, наблюдается ли перенос навыков запоминания пар «лицо—имя» на другие парные ассоциации или на ситуации запоминания вне лаборатории? (Ответ, кстати, неоднозначен, так как исследование не предусматривало оценку ни одной из этих форм переноса.) Второй общий момент — это *стойкость изменений*: сохраняется ли достигнутое улучшение и после исследования или исчезает по окончании программы? Здесь, как мы увидели, Шейх и коллеги получили некоторые позитивные данные о том, что спустя 6 месяцев эффект все еще был заметен. Однако в подобного рода исследованиях отнюдь не всегда оценивается стойкость изменений. В действительности, многие проекты вмешательства — в частности, программы тренировки памяти — даже не предусматривают проверку переноса или стойкость изменений. А если такие проверки и предусматриваются, вывод из них однозначен: добиться краткосрочного эффекта легче, чем стойкого, а частных и специфических эффектов — легче, чем общих и неспецифических.

Последнее замечание касается масштабов вмешательства. В принципе, предполагается, что в целом, «чем больше, тем лучше» (больше времени действия программы вмешательства, больше разнообразных компонентов программы, больше изменений в жизни испытуемых и т. д.); и в основном эти ожидания оправданны, опять-таки не только в отношении тренировок памяти, но и в отношении ряда других программ вмешательства, направленных на другие популяции. Тут, однако, в

силу вступает принцип осуществимости. Ресурсы всегда ограничены, и часто просто невозможно сделать все необходимое для достижения улучшения у людей, которым мы пытаемся помочь. Кроме того, чем больше компонентов мы включаем в программу, тем сложнее определить, какие именно компоненты сыграли решающую роль в успехе. Это проблема *конструктивной валидности*: точного теоретического объяснения любых обнаруженных эффектов. Как правило, конструктивная валидность важна для вопросов теории, однако здесь она имеет и практический смысл. Целью многих программ вмешательства является не просто улучшение жизни испытуемых, для которых они разработаны, но и получение информации, дающей возможность психологам помочь изучаемым популяциям в целом. Поэтому вмешательство, как и оценка, имеет двойную цель: принести конкретную пользу изучаемой группе и потенциальную пользу популяции, входящей в группу риска. А для достижения этой второй цели нам необходимы программы, которые можно осуществить и широко применять, элементы которых понятны и поэтому могут точно выполняться разными специалистами в разных условиях.

Несколько глав книги Фишер и Лернер (Fisher & Lerner, 1994a) посвящены более детальному рассмотрению отдельных программ вмешательства и общих принципов исследования их эффективности. Две книги, написанные Зиглером и его коллегами (Zigler & Muenchow, 1992; Zigler & Styfco, 1993), послужат ценным источником информации об одном из наиболее известных и значимых проектов вмешательства в детском возрасте — проекте *Head Start*. Проект *Head Start*, в свою очередь, может служить материалом для обсуждения не только вмешательства, но и другой важной роли прикладных исследований: их вклада в социальную государственную политику. Как узнает читатель этих книг, *Head Start* — несмотря на многие недоразумения, неудачи и разочарования — остается наиболее ценным из того, что психология развития сделала для общества.

Резюме

Первый раздел этой главы посвящен условиям, в которых проводятся исследования в области психологии развития. Он построен на разграничении структурированных лабораторных условий и естественных полевых условий. Независимые переменные можно подвергать манипуляциям, а зависимые переменные — измерять в любых из этих условий, что дает четыре варианта плана исследования.

В исследованиях типа «лаборатория—лаборатория» максимальны и контроль независимой переменной, и точность измерения зависимой переменной. Поэтому, как правило, такие исследования обладают высокой внутренней валидностью. Кроме того, некоторые проблемы психологии развития (например, многие вопросы перцептивного развития) можно изучать исключительно в контролируемых условиях лаборатории. С другой стороны, проводимые в лаборатории манипуляции и измерения зачастую резко отличаются от интересующих исследователя ситуаций, возникающих в реальной жизни; и эта искусственность ставит под угрозу внешнюю валидность. Кроме того, в лабораторных исследованиях особенно высок риск реактивности и искажения ответов.

Полевые исследования имеют сильные и слабые стороны, абсолютно противоположные сильным и слабым сторонам лабораторного исследования. Большим преимуществом исследования типа «полевые условия — полевые условия» является то, что и независимая, и зависимая переменная находятся здесь в естественной среде. Поэтому снижается проблема искусственности, следствие которой — неудовлетворительная внешняя валидность. Главное ограничение данного подхода касается осуществимости — контроль независимой переменной и измерение зависимой переменной в естественных условиях осуществить сложнее, чем в лабораторных, а с некоторыми переменными такой контроль и такие измерения вообще невозможны.

Плюсы и минусы комбинированных планов («лаборатория — полевые условия» и «полевые условия — лаборатория») можно вывести из характерных особенностей классических лабораторных и полевых исследований. Кроме того, их преимущество заключается во временном и пространственном разделении независимой и зависимой переменных, что иногда повышает валидность.

За обзором исследовательских планов идет их анализ и оценка. Обсуждаются сложности в классификации исследовательских подходов, затем рассматриваются показатели частоты использования разных вариантов исследования по данным обзора журнальных статей. Этот обзор свидетельствует об обоснованности критики исследователей в области психологии развития за их ориентированность на работу преимущественно в лабораторных условиях. В данном разделе подчеркивается значение *конвергентных операций* — изучения определенной проблемы максимальным количеством способов с целью компенсации неизбежных ограничений каждого из подходов достоинствами других подходов.

В среднем разделе главы рассматривается контекст исследования, но уже с несколькими иными позиций. В своей теории *экологических систем* Бронфенбреннер выделяет четыре взаимосвязанных слоя средового контекста, в котором происходит развитие. *Микросистема* — слой, ближайший к ребенку; он состоит из социальных агентов (например, родителей, сверстников) и условий, непосредственно влияющих на развитие ребенка. Под термином *мезосистема* подразумеваются отношения между микросистемами, к примеру влияние родителей на школьные успехи ребенка. Термином *экзосистема* Бронфенбреннер обозначает социальные системы, которые могут воздействовать на ребенка только опосредованно, например, попечительские школьные советы, местные власти, предприятия, где работают родители. Наконец, *макросистема* — это культура или субкультура, в которой развивается ребенок.

Общим выводом из теории Бронфенбреннера является необходимость изучения средовых контекстов, выходящих за пределы микросистемы, на которой сосредотачивается большинство исследователей. Особый интерес представляет взаимодействие между двумя и более экологическими системами — например, зависимость эффектов родительского воспитания от типа культуры или субкультуры. Рассматриваются два примера исследований, в которых было продемонстрировано такое взаимодействие. Анализируются также результаты обзора журналов, свидетельствующие о том, что уровень микросистемы остается наиболее типичным объектом изучения в психологии развития.

Заключительный раздел этой главы посвящен целям исследования. Особое внимание уделяется *прикладному исследованию*, то есть исследованию, цель которого в предупреждении и коррекции нарушений в развитии, а также в оптимизации процесса развития. Рассматриваются две формы прикладного исследования. Цель *оценки* — выявление значимых, чаще всего негативных, характеристик некой популяции, входящей в группу риска. В качестве примера оценочной методики приводится Оценочная шкала поведения новорожденного (ОШПН). Одним из свидетельств валидности этой шкалы является возможность с ее помощью обнаружить эффекты факторов риска пренатального и натального периодов; еще одним свидетельством валидности является возможность спрогнозировать статус ребенка. Однако прогностические возможности этой методики далеки от идеальных — что в той или иной степени справедливо по отношению ко всем психологическим оценочным методикам.

Оценка нередко является этапом, предвещающим вторую форму прикладного исследования — *вмешательство* с целью коррекции или предупреждения нарушений. В качестве примера, иллюстрирующего особенности программ вмешательства, описаны программы тренировки памяти для пожилых людей. Одной из особенностей программ вмешательства является необходимость выявления потенциальных участников, задача, для решения которой важна точность оценочной методики. Оценка имеет значение и по окончании программы для определения ее эффективности. Актуальным вопросом является также проблема *переноса* приобретенных навыков на ситуации, выходящие за пределы контекста вмешательства, а также *стойкость достигнутых изменений*. Наконец, цель многих проектов вмешательства — разработка программ, которые могли бы получить широкое распространение; поэтому важно, чтобы исследовательские проекты были осуществимы и их основные элементы достаточно понятны.

Упражнения

Большинству из нас рассуждать о влиянии на развитие микросистемных факторов (родителей, друзей, учителей и т. д.) значительно легче, чем о влиянии контекстов другого уровня, которые рассматривал Бронфенбреннер в своей теории экологических систем. Задумайтесь о роли мезосистемы, экзосистемы и макросистемы в своем собственном развитии. Как эти контекстуальные факторы более высокого уровня повлияли на то, кем вы стали? Как иначе могло бы происходить ваше развитие, если бы иным оказался экологический контекст?

Проект *Head Start* работает в более чем 2000 центрах во всех 50 штатах. Посетите по крайней мере один или, если есть возможность, несколько центров *Head Start* в своем районе. Попытайтесь посетить их несколько раз, чтобы застать разные занятия. Ознакомившись с тем, как работает *Head Start*, подумайте, как можно оценивать эффективность таких программ. Какие оценочные методики, по вашему мнению, было бы необходимо использовать? Найдите книги по этой теме, на которые делались ссылки в этой главе, и сравните собственные идеи с тем, что действительно использовалось и используется в исследованиях, посвященных *Head Start*.

Глава 7

Статистика

Все психологические исследования включают статистическую обработку данных. Нельзя стать исследователем в области психологии, не обладая знаниями основ статистического анализа. Кроме того, не имея элементарных представлений о том, что лежит в основе разных статистических показателей, невозможно понять и оценить исследования других людей.

Сказанное выше хорошо известно всем студентам, достаточно далеко продвинувшимся в изучении психологической науки. Все они также знают, что овладение статистикой — долгий процесс, который требует прослушивания нескольких курсов и прочтения нескольких увесистых учебников. Я не буду пытаться ужать содержание нескольких книг до одной главы. Моя задача намного скромнее: изложить в качестве дополнения к материалам о тестах некоторые основы и принципы, стоящие за статистическими процедурами. Эта глава может быть введением или дополнением к курсу статистики, а также может служить материалом, помогающим освежить уже имеющиеся знания.

Изложение в этой главе идет в направлении от сравнительно простого к более сложному. Начнем с того, что вспомним цели использования статистических показателей, затем, в качестве примера статистического объяснения, приведем уже знакомый вам *t*-критерий. Остальная часть главы в значительной мере будет посвящена усложненным вариантам простого анализа с использованием *t*-критерия — разнообразным ситуациям, когда нужны другие статистические показатели, их типам, а также основаниям для выбора между этими показателями. Поскольку каждый случай рассматривается кратко и в общих чертах, для каждой из затронутых тем предлагается список литературы.

Использование статистических показателей

Используя статистические показатели, психологи преследуют две цели: описать данные и прийти к логическому заключению об их значении. Первая из этих целей проста и понятна, достижение второй цели связано с определенными трудностями.

Таблица 7.1

**Количество агрессивных действий в выборке детей дошкольного возраста
(гипотетические данные)**

Трехлетние мальчики	Трехлетние девочки	Четырехлетние мальчики	Четырехлетние девочки
5	0	2	3
4	0	27	3
0	10	3	0
14	3	34	10
5	0	3	1
15	18	38	4
0	2	0	3
2	0	4	0
9	5	19	4
5	15	35	3
3	0	3	5
2	6	2	11
1	0	3	0
16	10	18	1
3	6	10	3

Дескриптивные статистические показатели

Вернемся к одному из предыдущих примеров: исследованию агрессии методом наблюдения в условиях детского сада. Допустим, исследователь собрал данные, представленные в табл. 7.1. Как можно заметить, имеют место значительные индивидуальные различия в частоте агрессивных действий; вполне вероятно, что есть также различия между полами и возрастными группами. Но как разобраться в этом хаосе цифр и определить истинное положение вещей?

Первый шаг — охарактеризовать данные при помощи ряда дескриптивных статистических показателей. Большинство дескриптивных статистических показателей используется для выявления центральной тенденции, или преобладающей формы ответов в выборке. Чаще всего в качестве меры центральной тенденции выступает среднее арифметическое, или просто среднее. Из табл. 7.2 явствует, что группы испытуемых из нашего гипотетического исследования действительно имеют разный средний уровень агрессии.

Чаще всего среднее — наиболее информативный дескриптивный статистический показатель. Однако он не единственный, в ряде случаев знание лишь среднего не дает полного представления о полученных результатах. Сравним группы 3- и 4-летних мальчиков. Как явствует из табл. 7.2, средний уровень агрессии у старших мальчиков выше. Однако сырые данные из табл. 7.1 свидетельствуют о том, что фактически большинство показателей агрессии обеих групп достаточно близки. Более высокое среднее у старших детей явилось следствием наличия нескольких очень высоких показателей. Или же сравним группы 3-летних мальчиков и девочек. Полагаясь на средние значения из табл. 7.2, мы могли бы заключить, что эти группы имеют одинако-

вый уровень агрессии. Однако сырые данные из табл. 7.1 свидетельствуют о том, что эти средние значения имеют разные основания.

Приведенные выше примеры демонстрируют необходимость иных дескриптивных статистических показателей помимо среднего арифметического. Есть еще две меры центральной тенденции. Одна из них — медиана. Медиана — это центр распределения, выше которого находится одна половина показателей, а ниже — другая. Сравним вновь результаты 3- и 4-летних детей. Из табл. 7.2 явствует, что данные результаты имеют общую медиану — 4. Это свидетельствует о фундаментальном сходстве двух распределений, сходстве, которое мешает заметить разницу средних. В целом, медиана приобретает особое значение тогда, когда распределение асимметрично, то есть включает несколько необычно высоких или низких показателей. В таких случаях среднее может дать искаженную картину типичных ответов.

Таблица 7.2

Дескриптивные статистические показатели для данных из таблицы 7.1

	Среднее	Медиана	Мода	Стандартное отклонение
3-летние мальчики	5,6	4	5	5,38
3-летние девочки	5,0	3	0	5,88
4-летние мальчики	13,4	4	3	13,90
4-летние девочки	3,4	3	3	3,29
Мальчики в целом	9,5	4	3	11,09
Девочки в целом	4,2	3	0	4,75
3-летние в целом	5,3	3,5	0	5,55
4-летние в целом	8,4	3	3	11,15

Третья мера центральной тенденции — мода. Мода — показатель, наиболее часто встречающийся в определенной группе. Эта мера используется редко, однако в некоторых обстоятельствах ее значение довольно информативно. Рассмотрим, к примеру, данные 3-летних девочек из табл. 7.1. Ранее мы отметили, что средний уровень агрессивных действий в этой группе — 5,0 — практически такой же, как и у мальчиков. Однако, в отличие от мальчиков, для 3-летних девочек модальным было нулевое значение. Этот факт вполне заслуживает того, чтобы упомянуть о нем в отчете.

Наряду с центральной тенденцией, дескриптивные статистические показатели характеризуют изменчивость распределения. Нам необходимо знать не только, какова центральная тенденция, но и то, насколько приближаются показатели к центральному значению или отклоняются от него. Чаще всего мерой изменчивости слу-

жит дисперсия. При ее расчете сначала находят среднее для выборки. Затем определяется разница между этим средним арифметическим и показателем каждого из испытуемых. Эти значения разности, или «отклонения», возводятся в квадрат, суммируются, а полученная сумма делится на $N - 1$, результатом чего и является показатель дисперсии. Таким образом, дисперсия — это приблизительно среднее квадратичных отклонений; «приблизительно», поскольку делитель равен $N - 1$, а не N . Чем больше разница между индивидуальными показателями, тем больше дисперсия.

В научных статьях в качестве меры изменчивости обычно указывается не дисперсия, а стандартное отклонение. Стандартное отклонение — это просто квадратный корень из показателя дисперсии. В табл. 7.2 он подсчитан для каждой из групп нашего гипотетического исследования. Полученные значения стандартного отклонения подтверждают наши интуитивные предположения о степени разброса индивидуальных показателей в группах. Обратите особое внимание на весьма значительное стандартное отклонение у 4-летних мальчиков, в группе, где было отмечено несколько крайне высоких показателей.

Статистические показатели, выводимые логическим путем

Предположим, мы получили значения среднего арифметического, представленные в табл. 7.2. Оказывается, что уровень агрессии изменяется как функция от возраста и пола. Но как выяснить наверняка, является ли обнаруженное различие истинным или это просто случайные колебания? На этот вопрос призваны ответить статистические показатели, выводимые логическим путем.

Для объяснения смысла статистических показателей, выводимых логическим путем, нужно вспомнить некоторые разграничения (имеющие частичное совпадение), введенные в предыдущих главах. Одно из них — разграничение между истинными показателями и погрешностями измерения. Любой показатель состоит из двух компонентов: действительного результата испытуемого, полученного при измерении, и любого рода погрешности измерения, возникающей при попытке выявить этот истинный показатель. Второе разграничение — между первичной дисперсией и вторичной дисперсией, или дисперсией ошибки. Первичная дисперсия связана с изучаемыми независимыми переменными; вторичная дисперсия, или дисперсия ошибки, обусловлена действием всех других факторов, то есть может иметь какой угодно источник, за исключением независимых переменных. Последнее разграничение — между популяцией и выборкой. Популяция — это весь тот контингент людей, который интересует исследователя; а выборка — это группа людей, фактически включенных в исследование.

При сравнении двух выборок (двух возрастов, двух полов, экспериментальных условий и т. д.) нас интересует вопрос, есть ли истинное различие между популяциями, из которых отобраны эти группы. Если бы нам удалось собрать данные по всей популяции, а не только по выборке, и исключить возможность погрешности измерения, у нас был бы ответ: полученные результаты и были бы результатами интересующей нас популяции. Однако, разумеется, сделать этого мы не можем; выборки — это всегда лишь часть популяции, измерение всегда неточно, и всегда существуют посторонние источники дисперсии. Именно поэтому нам необходимы методы оценки, или определения на основе логических заключений вероятности того, что выявленные различия между выборками отражают истинные различия между популяциями.

Поясним сказанное выше на примере гипотетического исследования агрессии и вопроса различий между полами в уровне агрессии. Мы уже знаем, что различия между полами действительно есть, в том смысле, что показатели мальчиков и девочек неодинаковы. Однако мы знаем и то, что это различие может объясняться погрешностями измерения и побочными источниками дисперсии. Кроме того, мы наблюдали лишь небольшую выборку из популяции, которая нас интересует — только 60 детей из миллионов 3- и 4-летних, посещающих детские сады США, и только несколько часов из жизни этих детей. Возможно, понаблюдав за теми же детьми вновь, мы получили бы несколько отличные результаты. Возможно, что, понаблюдав вторую выборку из 60 детей, мы опять-таки получили бы иные результаты. И возможно, что если бы нам удалось понаблюдать всю интересующую нас популяцию, мы получили бы еще какую-то совокупность данных. Именно для определения вероятности всех этих «возможно» необходимы статистические показатели, выводимые логическим путем.

В предыдущем абзаце цели использования статистических показателей, выводимых логическим путем, рассматриваются с двух позиций. Во-первых, с точки зрения воспроизводимости результатов или надежности. Получим ли мы одинаковые результаты, вновь и вновь производя один и тот же эксперимент? Во-вторых (что в действительности то же самое), с точки зрения перехода от выборки к популяции. Достаточно ли велико отличие, обнаруженное в выборке, чтобы доказать существование отличия в популяции? Как бы мы ни формулировали вопрос, нужно выбрать одно из двух: либо наши результаты действительно отражают положение вещей в популяции, либо они обусловлены действием случайных факторов, действующих в нашем конкретном исследовании. И как бы мы ни формулировали вопрос, использование статистических показателей, выводимых логическим путем, не дает однозначного ответа о том, что из сказанного верно; все, о чем мы можем судить по этим статистическим показателям — это о вероятности каждой из альтернатив. Это, фактически, главное, что нужно осознать в отношении статистических выводов: они вероятностны, а не абсолютны.

Теперь обратимся к конкретному примеру статистического анализа. Рассмотрим вновь различия между полами в уровне агрессии. Нам нужно определить, отражает ли обнаруженное в исследовании различие истинное различие в популяции или же оно — результат случайности. Как уже отмечалось, в качестве примера, за основу мы возьмем логику статистического анализа при использовании t -критерия.

Формула расчета t -критерия представлена ниже. Логика, положенная в основу этой аналитической проверки, довольно проста. Величина t -критерия, а следовательно, и вероятность того, что результаты неслучайны, зависит от трех факторов. Первый — разница между значениями средних. Чем больше различие, тем больше t . Второй — изменчивость внутри сравниваемых групп. Именно изменчивость представлена в довольно громоздком делителе. Чем она меньше, тем больше t . Наконец, третий фактор — объем выборки. Объем выборки влияет на конечный результат по двум направлениям. Во-первых, как можно заметить, проанализировав формулу, объем выборки влияет на изменчивость: чем больше n , тем меньше делитель. Во-вторых, даже подсчитав t , мы все еще должны определить,

какова вероятность того, что это значение t — результат случайности. Эта вероятность зависит как от величины t , так и от объема выборки. Чем больше n , тем ниже вероятность того, что полученное значение t — всего лишь результат случайных колебаний.

$$t = \frac{M_1 - M_2}{\sqrt{\frac{\sum x_1^2 - \frac{(\sum x_1)^2}{n_1} + \sum x_2^2 - \frac{(\sum x_2)^2}{n_2}}{(n_1 + n_2 - 2)}} * \sqrt{\frac{n_1 + n_2}{(n_1) * (n_2)}}}$$

M_1 — среднее арифметическое для группы 1;

M_2 — среднее арифметическое для группы 2;

n_1 — объем выборки для группы 1;

n_2 — объем выборки для группы 2;

x_1 — индивидуальные показатели испытуемых из группы 1;

x_2 — индивидуальные показатели испытуемых из группы 2.

Применим теперь формулу t для оценки различий между мальчиками и девочками в нашем гипотетическом исследовании. Получаем t , равное 2,41. Обратившись теперь к таблице показателей t -критерия (которая есть в любом учебнике по статистике), мы устанавливаем, что такое или большее значение t могло быть результатом случайности менее чем в 5 случаях из 100. Этот расчет вероятности случайности основан на том, что называется *нуль-гипотезой* — то есть предположении, что между группами в действительности нет различий. Результаты, вероятность случайного появления которых составляет менее 5 %, условно считаются статистически значимыми. Поэтому мы можем отвергнуть нуль-гипотезу об отсутствии различий между полами и заключить, что мальчики действительно агрессивнее девочек.

Мы еще вернемся к понятию статистической значимости. Однако сначала стоит повторить логику анализа с использованием t -критерия, поскольку она применима в отношении ряда других статистических критериев. Как отмечалось, эта логика действительно довольно очевидна и сводится к трем правилам, основанным на простом здравом смысле:

1. Случайность больших различий между группами менее вероятна, чем случайность небольших различий. Поэтому разница между большинством других значений среднего арифметического из табл. 7.2 (например, между 3-летними девочками и 3-летними мальчиками) слишком мала, чтобы дать существенный t -показатель, и поэтому, скорее всего, объясняется случайностью.
2. Чем меньше внутригрупповая изменчивость, тем меньше вероятность того, что различия являются результатом случайности. Небольшое число существенных отклонений от группового среднего в ту или иную сторону практически не отражается на значении среднего арифметического. Этот фактор играет роль при сравнении результатов 3- и 4-летних мальчиков. Несмотря на существенную разницу между средними показателями, сравнение с использованием t -критерия

говорит об отсутствии значимых различий, в немалой степени из-за высокой изменчивости в группе 4-летних мальчиков.

3. Наконец, вероятность случайности различий, обнаруженных в больших выборках, меньше, чем вероятность случайности таких же различий в небольших выборках. Если количество испытуемых невелико, один или два крайне высоких или крайне низких показателя могут исказить среднее арифметическое; в больших выборках такие случайные колебания компенсируются. Этот фактор играет роль при сравнении результатов 3-летних и 4-летних мальчиков. Если бы объем выборки составлял 30, а не 15 человек на группу, то полученный показатель t был бы значим.

Из сказанного выше следует, что цель использования статистических процедур, основанных на логических построениях, — установление статистической значимости. Важно ясно представлять, что подразумевается, а также, что *не* подразумевается под выражением «статистическая значимость».

Вспомним для начала, что выводы, основанные на логических статистических показателях, носят вероятностный характер. Утверждение, что определенное различие средних статистически значимо, означает, что такое различие *вероятнее всего* не случайно, если исходить из нуль-гипотезы об отсутствии различий в популяции. Однако всегда есть вероятность ошибки. Первая состоит в ошибочном отвержении нуль-гипотезы, то есть в выводе о наличии некоего эффекта при его реальном отсутствии. Этот тип ошибки называют ошибкой первого рода. В нашем исследовании агрессии мы допустили бы ошибку первого рода, заключив, что мальчики и девочки различаются по уровню агрессии, в то время как в действительности на уровне популяции в целом такое различие отсутствует. Вероятность ошибки первого рода определяется уровнем вероятности, на котором мы отвергаем нуль-гипотезу. Если уровень вероятности 0,05, риск допустить ошибку первого рода составляет 5 из 100. Если уровень вероятности ниже, скажем, 0,01 или 0,001, тогда, естественно, у нас гораздо меньше шансов ошибиться.

Второй тип ошибки состоит в принятии нуль-гипотезы тогда, когда в действительности имеется истинный эффект. Этот тип ошибки называется ошибкой второго рода. В исследовании агрессии мы допустили бы ошибку второго рода, если бы группы 3-летних и 4-летних детей различались, но мы заключили бы, что между ними нет различий. Вероятность ошибки второго рода рассчитать труднее, чем вероятность ошибки первого рода, и здесь мы даже не будем пытаться объяснить этот расчет. Однако замечу, что вероятность одной ошибки находится в обратной зависимости от вероятности второй ошибки, то есть чем выше вероятность одной, тем ниже вероятность другой. Исследователь, к примеру, может снизить риск ошибки первого рода, установив уровень вероятности 0,01, однако в то же время он существенно повышает риск ошибки второго рода. Отметим также, что психологи предпочитают минимизировать вероятность ошибки первого рода. Эта осторожность в позитивных выводах отражена в общепринятой норме: «значимыми» признаются только результаты, вероятность случайности которых составляет менее 5 %¹.

¹ Розноу и Розенталь (Rosnow & Rosenthal, 1989) критикуют позицию исследователей, целиком полагающихся на уровень вероятности 0,05: «Несомненно, Господь любит 0,06 ничуть не меньше, чем 0,05» (p. 1277).

Рассмотрение ошибок первого и второго рода возвращает нас к понятию валидности. В главе 2 рассказывалось о трех из четырех основных форм валидности. Четвертая форма — валидность статистического вывода: точность статистического вывода, сделанного при анализе данных. Верны ли наши заключения о наличии или отсутствии связи между переменными? Избежав ошибочного вывода о существовании связи при ее отсутствии (ошибка первого рода) и об отсутствии связи при ее наличии (ошибка второго рода), мы достигаем валидности статистического вывода.

Установив статистическую значимость, мы можем сказать, что наши результаты, вероятнее всего, не случайны. Важно отдавать себе отчет, что критерий значимости имеет отношение *только* к возможности случайных результатов. Значимость не исключает возможности искажения валидности. По этому критерию можно судить о наличии различий между двумя группами, но не о причинах различий.

Рассмотрим различие между полами в нашем исследовании агрессии. Нас интересует вероятность того, что это различие в поведении истинно (однако, разумеется, причины его еще предстоит выявить). Но значимое различие вполне могло появиться и по другим причинам. Возможно, наши наблюдатели ожидали от мальчиков или от девочек большей агрессивности и поэтому в соответствующем направлении искажали результаты — отсюда различие, обусловленное необъективностью наблюдателей. Возможно, на девочках сильнее отражается присутствие наблюдателя, и поэтому они более склонны подавлять агрессию, когда за ними наблюдают, — отсюда различие, обусловленное дифференцированной реактивностью. Возможно, мы наблюдали девочек в начале года, а мальчиков позже, когда агрессия становится обычным явлением, — отсюда различие, обусловленное одновременным влиянием фактора принадлежности к определенной группе и фактора времени измерения.

Суть в том, что любые из описанных в этой книге факторов, ставящих под сомнение валидность, могут все еще действовать, искажая наши результаты. Статистическая значимость не гарантирует общей валидности. Это лишь отправная точка, необходимое, но не достаточное условие для вывода о том, что мы обнаружили что-то существенное.

И последнее замечание — статистическая значимость не гарантирует того, что результаты имеют некоторую научную ценность. «Значимость» в том смысле, в котором этот термин употребляется здесь, имеет отношение только к статистической вероятности, а не к теоретической или практической важности. Различие между полами в уровне агрессии может быть истинным, в том смысле, что оно не случайно и не обусловлено неудовлетворительной валидностью. Достаточно ли велико это различие, чтобы что-то значить — в отношении, например, того, как воспитатели должны вести себя с мальчиками и девочками, — отдельный вопрос. Важно помнить, что статистическая значимость различия зависит не только от величины разницы, но и от объема выборки. В достаточно большой выборке даже незначительное различие достигает уровня значимости. Мы еще вернемся к этому вопросу, когда будем рассматривать величину эффекта.

Выбор статистического показателя

Для многих студентов слово «статистика» ассоциируется с зазубриванием формул и бесконечными часами утомительных подсчетов. В действительности профессиональный исследователь если и может воспроизвести, то не более нескольких формул и тратит на расчеты совсем немного времени. В этом нет необходимости: формулы есть в учебниках или заложены в компьютер, а расчеты можно производить на калькуляторе, на компьютере (или предоставить это студенту-лаборанту!). Что значительно важнее, это знать, какого рода статистический анализ подходит и информативен для определенного рода данных. При выборе наиболее подходящего статистического показателя учитывается множество факторов. В этом разделе мы рассмотрим три из них: уровень, на котором измеряется зависимая переменная, распределение значений зависимой переменной и план исследования.

Уровень измерения

Понятие уровня, или шкалы измерения было введено в главе 4. Вспомним, что выделяют четыре уровня измерения: *номинальный*, или качественное обозначение результатов; *порядковый*, или ранжирование результатов по некой шкале количественных значений; *интервальный*, или распределение результатов по шкале количественных значений, которые не только упорядочены, но и равноудалены друг от друга; и уровень *отношений*, или равномерное упорядочение результатов по шкале количественных значений, имеющей абсолютный нуль.

Уровень измерения является одним из факторов, определяющих, какой из статистических критериев уместнее всего употребить. Некоторые критерии, включая и t , используются только тогда, когда измерение производится на шкале интервалов или шкале отношений. Основание для этого требования станет очевидным при анализе формулы на рис. 7.1. Для расчета t -критерия мы должны произвести ряд арифметических операций с числами — сложить, а затем разделить, чтобы получить среднее, вычесть каждое число из среднего, чтобы получить показатель отклонения и т. д. Эти операции имеют смысл только в том случае, если числа, с которыми мы работаем, являются точным отображением количественного значения, а не просто названиями или порядковыми номерами. Показатели частоты из табл. 7.1 отвечают указанному требованию, и, следовательно, к этим данным t -критерий применим. Однако t -критерий не подошел бы, если бы наши данные были основаны на описанной ранее рейтинговой шкале. Мы могли бы, к примеру, сложить рейтинговую оценку 5 («крайне агрессивный») с рейтинговой оценкой 1 («совершенно неагрессивный») и получили бы среднее 3 («умеренно агрессивный»). (Вскоре я уточню это замечание. Кроме того, необходимо помнить, что не все специалисты в области теории измерения и статистики сходятся во мнении по вопросу связи между шкалами измерения и статистическими показателями, — см. Cliff, 1993; Michell, 1986.)

Распределение показателей

Использование некоторых статистических критериев связано с определенными предположениями о распределении оцениваемых этим критерием показателей. В частности, так называемые параметрические критерии зависят от определенных

предположений о распределении данных. Это, фактически, и является смыслом термина «параметрический»: статистический анализ зависит от валидности некоторых предположений в отношении «параметров» популяции, к которой принадлежит выборка. Рассмотренный выше t -критерий — пример параметрического критерия; критерий, используемый в дисперсионном анализе ($ANOVA$), которому посвящен следующий раздел, — еще один пример.

Если говорить более конкретно, в основе использования большинства параметрических критериев лежит два допущения. Первое состоит в том, что показатели распределены по закону нормального распределения. Второе — что дисперсия в сравниваемых группах одинакова. Второе допущение распространяется не на все случаи, но применимо ко многим, часто используемым параметрическим критериям, включая t -критерий и F -критерий дисперсионного анализа.

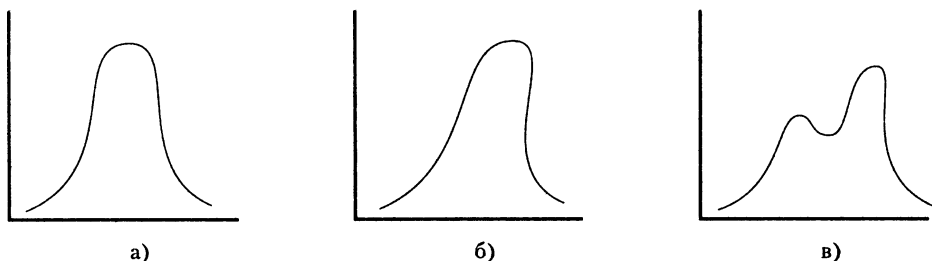


Рис. 7.1. Примеры нормального и ненормального распределения

Мы уже обсуждали понятие дисперсии. Рассмотрим теперь необходимые условия нормального распределения. На рис. 7.1 (а) изображено нормальное распределение. Термин «нормальное распределение» используется в отношении классической колоколообразной кривой, к распределению, в котором среднее, медиана и мода совпадают, а показатели постепенно уменьшаются по мере удаления от этого центра. Рис. 7.1 (б) и (в), напротив, иллюстрируют распределение, явно отличное от нормального.

Между уровнем измерения и распределением есть определенная связь. Показатели номинальных и порядковых шкал не могут иметь нормальное распределение. Что касается номинальной шкалы, в ней нет количественных значений, и поэтому вопрос распределения по шкале количественных значений не стоит; все, что здесь возможно, это подсчет частоты случаев в каждой из категорий. Если говорить о порядковой шкале, то нам неизвестна разница между показателями, а следовательно, и их распределение. Кроме того, в абсолютно упорядоченной шкале (то есть при отсутствии совпадений) на каждый уровень шкалы приходится всего по одному случаю; поэтому теоретически распределение будет плоским. Таким образом, необходимым условием нормального распределения является наличие шкалы отношений или интервалов. Тем не менее это недостаточное условие, поскольку кривая показателей все еще может выглядеть так, как на рис. 7.1 (б) или (в). Однако по закону нормального распределения могут быть распределены только показатели, соответствующие определенным шкалам.

Мы только что рассмотрели предположения, лежащие в основе использования параметрических критериев t и F . Скажем теперь несколько слов об альтернативе

Логика проверки с использованием критерия хи-квадрат: определение того, насколько полученные значения частоты в каждой клетке таблицы отклоняются от ожидаемой частоты, если допустить отсутствие различий между группами.

Фактическая частота	Игрушечные машинки	Конструктор	Куклы и игрушечная мебель	Принадлежности для рисования и рукоделия
Мальчики	12	10	1	5
Девочки	1	4	12	11
Ожидаемая частота				
Мальчики	6,5	7	6,5	8
Девочки	6,5	7	6,5	8

$$\text{Формула: } \chi^2 = \sum \frac{(\text{фактическая частота} - \text{ожидаемая частота})^2}{\text{ожидаемая частота}}$$

$$\chi^2 = 23,42; \text{ уровень вероятности } < 0,01$$

Вывод: между мальчиками и девочками существуют значимые различия в предпочтении игрушек.

Рис. 7.2. Иллюстрация аналитической процедуры с использованием критерия хи-квадрат

параметрических критериев, после чего можно будет сделать еще несколько замечаний относительно выбора статистического показателя.

Как можно было ожидать, альтернативой параметрическим являются **непараметрические критерии**. Рисунок 7.2 служит иллюстрацией для широко используемого непараметрического критерия хи-квадрат. Гипотетические данные, представленные на рисунке, относятся к описанному ранее исследованию предпочтений игрушек; гипотетический результат состоит в том, что предпочтение определенных игрушек является функцией от пола¹. Хи-квадрат используется при наличии номинальных данных, таких, как данные, представленные на рисунке, для которых *t*-критерий не подходит. Для каждого из четырех уровней измерения — номинального, порядкового, интервального и уровня отношений — существуют свои непараметрические критерии. Таким образом, этот подход имеет более широкое применение, чем использование параметрических критериев. Кроме того, непараметрические показатели не связаны с предположениями о виде распределения, которые лежат в основе параметрических показателей; поэтому непараметрические критерии применимы к данным, построенным на шкалах интервалов и отношений, но не удовлетворяющим параметрическим допущениям. (Из работ, посвященных непараметрическим критериям, можно назвать следующие: Gibbons, 1993; Marasculio & McSweeney, 1977; Siegel & Castellan, 1988.)

По какому принципу осуществляется выбор между параметрическими и непараметрическими характеристиками? Как только что отмечалось, в ряде случаев выбора просто нет, поскольку единственный вариант — это непараметрический

¹ Формулу, представленную на рисунке, называют «определяющей формулой» хи-квадрат. Есть также «калькуляционная формула» — равноценная в математическом смысле, но более удобная для проведения расчетов. Формулы многих статистических показателей также разделяются на определяющие и калькуляционные.

критерий. В других случаях необходимо принять решение, и здесь приобретает значение несколько понятий. Рассмотрим два из них: мощность и устойчивость.

Термин **мощность** означает вероятность того, что определенный логический критерий исключит нуль-гипотезу тогда, когда ее действительно нужно исключить. Чем мощнее критерий, тем лучше он выявляет истинные различия и поэтому позволяет безошибочно отвергнуть нуль-гипотезу. Это понятие, вероятно, кажется знакомым, поскольку мощность — это еще один способ охарактеризовать ошибку второго рода. Чем мощнее критерий, тем меньше вероятность ошибки второго рода.

В некоторых случаях параметрические критерии мощнее аналогичных непараметрических критериев. По сути, это объясняется тем, что при расчете параметрического критерия используется больше информации о данных. Многие непараметрические критерии, например, ограничены порядковыми характеристиками данных, в частности, рангом показателей в сравниваемых выборках. При расчете *t*-критерия, напротив, задействуются фактические показатели и абсолютная разница между ними; поэтому иногда с его помощью выявляются различия, которые не смогли выявить непараметрические критерии. Следует добавить, что разница в мощности, как правило, невелика и обнаруживается преимущественно при изучении больших выборок. Кроме того, она не является чем-то неизбежным. Во многих ситуациях параметрические и непараметрические критерии обладают одинаковой мощностью. Если предположения, лежащие в основе параметрического критерия, серьезно нарушаются, непараметрические аналоги могут оказаться более мощными (см. Blair & Higgins, 1980).

Сказанное о параметрических предположениях подводит нас к понятию устойчивости. **Устойчивость** характеризует безопасность отклонений от допущений, лежащих в основе некоего критерия. Устойчивый критерий сравнительно нечувствителен к таким нарушениям, то есть, как правило, он позволяет сделать точные выводы о значимости даже тогда, когда допущения не соответствуют действительности. И *t* и *F*-критерии достаточно устойчивы. Именно поэтому в литературе можно часто встретить указание на их использование даже для данных, не отвечающих рассмотренным выше требованиям — данным рейтинговых шкал, к примеру, или данных, распределение которых заметно отличается от нормального, или при наличии неравной дисперсии у сравниваемых групп. Устойчивость не означает, что исследователь может, не задумываясь, применять параметрические критерии к любому типу данных; однако не следует и слишком поспешно отказываться от параметрических критериев лишь потому, что некое допущение, лежащее в их основе, нарушается. Возможно, стоит посоветоваться со специалистом: применим ли выбранный параметрический показатель к имеющимся данным?

План исследования

Мы рассмотрели две детерминанты выбора статистического показателя: уровень измерения и распределение данных. Третий фактор, который следует учитывать, — это план исследования.

Имеют значение разные аспекты плана. Один из аспектов — количество уровней независимой переменной. В нашем примере с агрессией детей дошкольного возраста этот фактор довольно прост: две возрастные группы и два пола. Поэтому здесь достаточно легко при сравнении двух уровней каждой переменной можно

применить t -критерий. Предположим, однако, что мы усложняем ситуацию, добавляя дополнительные уровни. Поскольку с полом представить это себе довольно трудно, включим новые возрастные группы. Допустим, вместо двух у нас их шесть. Что происходит тогда с нашим t -критерием?

Наиболее очевидным следствием является то, что возникает необходимость подсчитать значительно большее количество критериев. При наличии шести возрастных групп возможно 15 парных сравнений. Поэтому, чтобы что-то обнаружить, придется подсчитать значение 15 t -критериев. Рассчитывать 15 показателей и указать их все в отчете, естественно, довольно неудобно. Однако более серьезный довод против этого имеет отношение к уровню вероятности. Нам нужно, чтобы этот уровень оставался неизменным, какой бы рубеж для значимости мы ни выбрали — к примеру, традиционные 0,05. Однако наличие множества t делает интерпретацию уровня вероятности весьма затруднительной. Получив 15 значений, каждое из которых находится на уровне 0,05, мы получаем вероятность того, что значимость по крайней мере одного из этих показателей носит случайный характер, равную 0,54.¹ Как же тогда интерпретировать любой статистически значимый результат?

Проблема, в действительности, даже еще сложнее. Вероятность 0,54 основывается на предположении, что все 15 показателей независимы друг от друга. Однако, как правило, это не так; они взаимосвязаны в том смысле, что одни и те же данные используются для разных сравнений. Это, фактически, относится к описанному выше случаю сравнения между возрастными группами: каждая из шести возрастных групп вносит свои данные — одни и те же — в расчеты 5 из 15 критериев. При наличии такого рода взаимозависимости критериев определить точный уровень вероятности для каждого критерия невозможно. Исследователь может подсчитать значение какого-то t и выявить значимость на уровне 0,05; однако вполне может оказаться, что истинный уровень значимости совершенно иной.

Есть и еще одна проблема, связанная с множественностью t . Допустим, что мы усложнили наше исследование не добавлением уровней независимой переменной, а введением дополнительных независимых переменных. Помимо возраста и пола как детерминант агрессии мы могли бы изучать эффект обстановки игровой комнаты, разницу между поведением в группе и на улице, влияние показа половине детей агрессивного мультлика и т. д. Ясно, что чем больше независимых переменных, тем большее количество t нужно подсчитать. Но проблема состоит не только в избытке показателей t . При изучении множественных переменных всегда существует вероятность зависимости эффекта одной переменной от уровня другой. Иными словами, возможно взаимодействие переменных. Эффекты взаимодействия необходимо выявить, но это довольно трудно сделать, используя только t -критерий.

Чаще всего в качестве альтернативы использования t -критерия проводят дисперсионный анализ (ДА). По существу, ДА расширяет возможности t -критерия на те случаи, когда имеется более двух средних. Метод расчета здесь иной и более

¹ Возможно, проще всего увидеть то, откуда появляется такая вероятность, это задаться вопросом, каковы шансы не получить случайный результат. При применении одной статистической проверки вероятность избежать такой ошибки составляет 0,95. При проведении двух отдельных проверок вероятность избежать ошибки определяется значением двух значений вероятности, то есть 0,952. При проведении 15 проверок эта вероятность составит 0,9515 или 0,46. Поэтому вероятность того, что мы получим хотя бы один случайный значимый результат, равна $1 - 0,46$.

сложный, чем метод расчета t , и в этой книге мы даже не будем пытаться его описать. Однако логика, лежащая в основе обоих приемов, одинакова: мы проверяем значимость, определяя, насколько первичная дисперсия, связанная со сравниваемыми группами, превышает вторичную дисперсию или дисперсию ошибки. Статистический показатель, являющийся результатом этой проверки, обозначается буквой F , и значимость его, как и значимость t , устанавливается по стандартным таблицам, которые можно найти в любом учебнике по статистике.

Рассмотрим, как можно было бы применить ДА в исследовании агрессии. У нас две независимые переменные: возраст и пол. Чтобы более наглядно продемонстрировать преимущества ДА перед использованием t , предположим, что в действительности, переменная возраста имеет 6 уровней, а не 2, как указано в табл. 7.1. Применение ДА дает показатель F для каждой из независимых переменных, или значения главных эффектов. Если для пола F значим, тогда в отношении этой переменной наши действия завершены; поскольку переменная пола имеет только два уровня, мы можем просто посмотреть на средние значения, чтобы определить, в чем состоит эффект. Значимый главный эффект для возраста — более сложный случай. Здесь показатель F основан на одновременном сравнении всех шести возрастных групп, а установление значимости подразумевает, что значим результат, по меньшей мере, одного парного сравнения. Тогда нам нужно использовать дополнительные критерии с тем, чтобы определить, результат какого из сравнений (или результаты каких сравнений) обладает значимостью. Эти дополнительные критерии сходны с t -критерием, однако их подсчет несколько проще и производится, только если общий показатель F свидетельствует о наличии значимого эффекта.

Результатом ДА является также третий показатель F — для взаимодействия между возрастом и полом. В целом, ДА дает столько F , сколько в исследовании существует возможных комбинаций независимых переменных. Если, к примеру, в исследовании 3 независимые переменные, результатом ДА будет четыре F , указывающих на взаимодействие: по одному на каждое парное сочетание и одно для тройного сочетания. Как и в случае со значимым главным эффектом, значимость взаимодействий можно проверять при помощи специальных критериев для выяснения основы эффекта взаимодействия.

Для выбора статистического показателя важен еще один аспект исследовательского плана. До этого момента мы говорили в основном о межсубъектных планах — то есть случаях, когда данные каждого испытуемого попадают только в одну категорию сравниваемых условий или групп. Но, как мы знаем из главы 3, многие независимые переменные изучаются при помощи внутрисубъектных планов, в которых данные каждого испытуемого попадают в категорию данных по каждому из экспериментальных условий. Что происходит со статистическими показателями, когда каждый испытуемый представлен в каждом из условий?

Ответ довольно прост: мы переходим от межсубъектных критериев, которые рассматривались до этого момента, к соответствующим внутрисубъектным критериям. В действительности, для каждого межсубъектного критерия, о которых мы говорили выше, существует свой внутрисубъектный аналог. К примеру, есть внутрисубъектный t -критерий, а также внутрисубъектный дисперсионный анализ или дисперсионный анализ с повторными измерениями. Кроме того, существуют непараметрические критерии, подходящие для внутрисубъектных данных (напри-

Таблица 7.3

IQ и результаты по тесту достижения в выборке 5-классников

Испытуемый	IQ	Тест достижений
1	82	22
2	85	18
3	90	43
4	92	28
5	95	23
6	99	24
7	101	48
8	102	30
9	104	56
10	107	35
11	108	38
12	112	46
13	114	27
14	116	54
15	124	50
16	140	60

мер, критерий изменения Макнемара, использование которого заключается в измерении нескольких хи-квадратов). Логика этих статистических процедур сходна с логикой использования межсубъектных критериев; однако в большинстве внутрисубъектных проверок анализируется действительная разница между показателями (например, результаты некоего испытуемого в условиях 1 минус его же результаты в условиях 2). Поскольку в центре внимания находятся показатели различия, данные критерии применимы не только для исследовательских планов с реальными повторными измерениями, но и для случаев, когда для каждого испытуемого в одних условиях подбирается соответствующий испытуемый в других условиях.

Следует сделать еще одно замечание, касающееся внутрисубъектных показателей. Оно повторяет то, что было сказано в главе 3 при обсуждении относительных преимуществ внутри- и межсубъектных планов. Тогда мы отметили, что внутрисубъектные критерии, как правило, обладают большей мощностью, чем аналогич-



Рис. 7.3. Диаграммы рассеяния, иллюстрирующие корреляции разного уровня

ные межсубъектные критерии. Это обусловлено уменьшением вторичной дисперсии, связанной с индивидуальными различиями испытуемых. Если в каждое из экспериментальных условий ставятся одни и те же испытуемые, вероятность внесения нежелательной дисперсии в результаты группового сравнения, обусловленные индивидуальными различиями, снижается. Большая мощность — одно из оснований для выбора между внутрисубъектными и межсубъектными подходами.

Меры связи

До этого момента основное внимание уделялось процедуре выявления различий между группами. Однако это не единственная область применения статистических процедур. Возьмем, к примеру, исследование, в котором были получены данные, представленные в табл. 7.3. Нас интересует вопрос, есть ли связь между *IQ* и успешностью выполнения стандартного теста достижения. Что нам нужно сделать?

Для данных из табл. 7.3 подходит **корреляционный статистический показатель**. Корреляция — это мера связи между двумя переменными. Как мы узнали из главы 3, значение корреляционного показателя находится в пределах от $+1$ до -1 . Коэффициент корреляции равный $+1$ свидетельствует о наличии абсолютно положительной связи между переменными, коэффициент корреляции равный 0 свидетельствует о полном отсутствии связи, а коэффициент корреляции равный -1 указывает на наличие абсолютно отрицательной связи. Эти варианты иллюстрируют графические изображения на рис. 7.3. Корреляционный показатель отличный от нуля свидетельствует о положительной или отрицательной связи, при этом сила связи увеличивается с приближением значения к $+1$ или -1 .

О чем же говорят данные, представленные в табл. 7.3? Для определения меры связи мы должны сначала выбрать соответствующий корреляционный показатель, поскольку для вычисления корреляции существует множество разных методов. Как и в случае с логическими критериями, выбор метода зависит от наших предположений относительно характера данных. Чаще всего используются два показателя: коэффициент корреляции произведения моментов Пирсона и коэффициент корреляции рангов Спирмена. Статистический показатель Пирсона — это параметрический критерий, использование которого основано на тех же допущениях, что и использование остальных параметрических критериев — а именно на допущении, что измерение происходило по шкале интервалов или отношений, а данные распределены по закону нормального распределения¹. Корреляционный показатель Спирмена — непараметрический критерий, основанный исключительно на порядковых характеристиках данных, и поэтому применяется чаще, чем критерий Пирсона. Оба показателя, надо заметить, зависят от другого важного предположения: что связь между переменными линейная. Если связь иного рода (к примеру, криволинейная, то есть при изменении значения одной переменной значение другой переменной сначала увеличивается, а затем уменьшается), стандартный корреляционный критерий неприменим.

¹ В действительности r сам по себе, как дескриптивный показатель, является непараметрическим, однако определение его статистической значимости зависит от параметрических предположений.

Порядок рангов			
<i>IQ</i>	Тест достижений	<i>IQ</i>	Тест достижений
1	2	9	15
2	1	10	8
3	10	11	9
4	6	12	11
5	3	13	5
6	4	14	14
7	12	15	13
8	7	16	16

$$\text{Формула: } r_{ho} = 1 - \frac{6 \sum D^2}{N(N^2 - 1)},$$

$r_{ho} = 0,70$, уровень вероятности $< 0,01$.

D — показатель отклонения, то есть абсолютная разница рангов;

N — количество пар.

Вывод: между *IQ* и результатами выполнения теста достижений существует значительная положительная корреляция.

Рис. 7.4. Логика коэффициента ранговой корреляции Спирмена: определение меры общности парных показателей в двух распределениях

Поскольку использование критерия Спирмена проиллюстрировать легче, применим его для данных табл. 7.3. Формула коэффициента Спирмена, а также применение ее к нашим данным представлены на рис. 7.4. Проанализировав формулу, можно заметить, что коэффициент Спирмена — это мера общности рангового порядка пар показателей двух распределений. Если согласованность рангов полная, тогда, при отсутствии показателей отклонения, вычитаемое будет равно нулю, а коэффициент корреляции +1. Чем чаще и сильнее показатели отличаются по рангу, тем дальше от единицы будет корреляционный показатель. В нашей выборке данных коэффициент корреляции между *IQ* и результатами теста достижений равен 0,7, что свидетельствует о достаточно тесной, но не абсолютной связи. Стоит отметить, что применение к этим данным корреляционного анализа Пирсона даст очень близкое значение: 0,71. Фактически для большинства данных значения коэффициентов Спирмена и Пирсона очень близки.

О чем говорит наличие корреляции между двумя переменными? Корреляция, как и среднее арифметическое или медиана, — дескриптивный статистический показатель, характеризующий, однако, не центральную тенденцию, а связь между переменными. Прежде чем интерпретировать значение коэффициента корреляции, необходимо проверить его статистическую значимость. Нуль-гипотеза при такой проверке заключается в том, что коэффициент корреляции между двумя переменными равен нулю; вопрос состоит тогда в том, есть ли значимое отклонение полученного коэффициента корреляции от нуля. Ответить на этот вопрос достаточно просто, поскольку в учебниках по статистике содержатся таблицы, по

которым непосредственно можно установить уровень вероятности для любых коэффициентов корреляции (многие компьютерные программы также осуществляют подсчет уровня вероятности). На значимость влияют и величина коэффициента корреляции, и объем выборки; с их повышением растет вероятность значимости. Из таблицы явствует, что коэффициент корреляции равный 0,7 в выборке объемом 16 (то есть при наличии 16 пар показателей) значим на уровне 0,01; таким образом, между IQ и уровнем достижений действительно имеется связь.

Значимость важна, но это лишь половина дела. Вспомним, что нас интересует не только существование связи, но и ее сила. Как правило, силу корреляционной связи интерпретируют с точки зрения точности прогноза: зная результаты испытуемого по одной переменной, насколько точно мы сможем предугадать его результаты по другой переменной? При корреляционном показателе, равном нулю, отношения между переменными носят случайный характер, и знание одного показателя не дает нам возможности предсказать другой показатель. По мере отклонения коэффициента корреляции от нуля его прогностическая способность возрастает, достигая максимума при коэффициенте корреляции равном ± 1 .

Другой (равноценный) способ рассмотрения корреляции — с точки зрения доли объясняемой дисперсии. Используя показатели по одной переменной для прогноза показателей по второй переменной, мы «объясняем», в статистическом смысле, определенную долю дисперсии значения второй переменной. Чем выше коэффициент корреляции, тем большая доля дисперсии получает объяснение. Эту закономерность можно определить точнее. Если коэффициент корреляции — пирсоновский r , доля объясняемой дисперсии составляет r^2 . Таким образом, коэффициент корреляции между IQ и уровнем достижений равный 0,71 означает, что вариации одного показателя объясняют 50 % вариаций другого.

Последнее из утверждений ограничивает нас в наших интерпретациях. Коэффициент корреляции равный 0,71 довольно высок, однако даже при таком значении половина дисперсии все еще не получает объяснения. С приближением коэффициента корреляции к нулю доля объясняемой корреляции уменьшается, и довольно стремительно. Коэффициент корреляции равный 0,5 объясняет 25 % дисперсии, а коэффициент корреляции равный 0,3 — лишь 9 % дисперсии.

Сказанное выше напоминает нам о разнице между значимостью и ценностью. Коэффициент корреляции может быть статистически значим, но в то же время столь мал, что его теоретическая или практическая ценность будет минимальна. Вероятность таких статистически значимых, но в действительности незначительных статистических показателей особенно велика при изучении больших выборок. В выборке объемом 50 коэффициент корреляции 0,27 достигает значимости на уровне 0,05. В выборке объемом 100 единиц значимостью обладает уже коэффициент корреляции 0,19.

Помимо объема выборки при оценке коэффициента корреляции важно учитывать диапазон значения переменных. Здесь могут возникнуть две проблемы. Чаще всего, это проблема, о которой говорилось в главе 46, сужение диапазона, которое происходит тогда, когда значения одной переменной так близки друг к другу, что разница между ними не связана с дисперсией значений других переменных. Предположим, сравнивая IQ и уровень достижений, мы решили ограничить выборку

детьми из классов для «одаренных». Как правило, отбор в эти классы производится по критерию IQ ; в нашем случае он соответствует, скажем, 130 пунктам и выше. Решение сосредоточиться исключительно на очень высоких IQ означает, что мы резко сужаем диапазон дисперсии значений одной из наших переменных; вместо 60–70 пунктов диапазон IQ будет составлять лишь около 20. При такой скученности показателей IQ разница между ними вряд ли обнаружит существенную связь с любым параметром, включая разницу в уровне достижений.

Возможно также, что диапазон значений переменной будет слишком широк. Допустим, от одного испытуемого к другому IQ изменяется на 20 пунктов: начиная с ребенка, имеющего $IQ = 40$, следующего ребенка с показателем $IQ = 60$ и т. д. до восьмого ребенка, обладающего $IQ = 180$. При столь большом разбросе велика вероятность того, что IQ будет значимо и существенно коррелировать практически со всеми показателями, которые мы измерим в нашей выборке. Сомнительно, однако, что величина таких коэффициентов корреляции будет иметь для нас большой смысл.

Слишком узок или слишком широк диапазон значений, в основе проблемы лежит недостаточная внешняя валидность. Для того чтобы коэффициент корреляции был для нас интересен, он должен характеризовать не только выборку, для которой был подсчитан, но и всю популяцию, которую представляет эта выборка. Поэтому выборка должна быть репрезентативна — как по параметрам центральной тенденции, так и по параметру диапазона дисперсии — в отношении популяции, частью которой она является. Если выборка будет нерепрезентативна, полученные при ее изучении коэффициенты корреляции не будут обладать достаточной внешней валидностью.

Альтернативы и более сложные варианты дисперсионного анализа

В этом разделе мы вернемся к вопросу сравнения групп. Не будет преувеличением сказать, что, по крайней мере последние 50 лет, дисперсионный анализ является самым распространенным методом такого сравнения. Однако, несмотря на свою популярность, ДА — не всегда предпочтительный вариант статистического анализа, даже в случаях, когда предположения, лежащие в его основе, справедливы. Есть ситуации, в которых более полное представление о полученных результатах дают другие формы анализа. В этом разделе мы кратко рассмотрим ряд таких ситуаций и соответствующие статистические процедуры. Еще менее чем в предыдущих разделах изложение претендует на полноту; цель заключается лишь в том, чтобы дать читателю возможность понять или переосмыслить разнообразные варианты более глубокого анализа и основания для его использования.

Плановые сравнения

Ранее мы отметили проблемы, возникающие при множественности значений t -критерия, следствием которых является необходимость использования дисперсионного анализа, позволяющего осуществить обобщенную проверку на наличие или отсутствие значимых эффектов. Однако обобщенные проверки не всегда облада-

ют преимуществом. ДА более всего подходит для тех случаев, когда у исследователя нет четких гипотез и ему скорее хочется выявить любые интересные результаты, которые может дать исследование. В подобных случаях такая комплексная проверка, как ДА, предпочтительнее множества одиночных проверок. Однако в некоторых исследованиях имеются четкие гипотезы, и основная цель проведения статистической проверки состоит в получении ясного ответа в отношении каждой из гипотез. В таких случаях ДА малоэффективен, поскольку эта форма анализа включает сравнения, которые могут и не интересовать исследователя, тем самым снижая эффективность сравнения того, что действительно интересно.

Рассмотрим пример (взятый из работы Haas, 1981). Нас интересует влияние обучения детей на их ответы при использовании нравственных дилемм Колберга (см. главу 12). Мы используем две формы обучения детей: наблюдение за взрослым, дающим ответы, которые соответствуют более высокому нравственному уровню, и обсуждение вопросов нравственности в группе сверстников. Нам также интересен возможный эффект сочетания двух типов обучения, поэтому мы включаем третье экспериментальное условие, при котором дети и наблюдают за отвечающим взрослым, и проводят обсуждение в группе. Мы знаем, что для оценки эффекта обучения необходима контрольная группа; поэтому четвертое условие состоит исключительно в претесте и посттесте, без каких-либо экспериментальных вмешательств. Однако нас беспокоит, что наблюдение образца поведения или внутригрупповое взаимодействие могут сами по себе повлиять на ответы, вне зависимости от нравственного развития, которое мы изучаем. Поэтому мы включаем еще два контрольных условия: одно — при котором дети наблюдают взрослого, рассуждающего о чем угодно, кроме нравственных проблем; и второе — при котором дети обсуждают любую тему, не касающуюся нравственности.

В таком исследовании нас в действительности не интересует главный эффект каждого из условий, эффект, который может быть обусловлен значимой разницей между любыми из возможных пар средних значений. Наш интерес более специфичен и ограничен сравнениями между средними значениями, которые важны для исследования. Мы, к примеру, захотим выяснить, отличается ли на уровне значимости каждое из экспериментальных условий от соответствующего ему контрольного, а также, различаются ли между собой три экспериментальных условия. Эти сравнения действительно имеют смысл. В других сравнениях гораздо меньше смысла — к примеру, бессмысленно сравнивать экспериментальное условие группового обсуждения и контрольное условие наблюдения поведения взрослого. Обобщенный дисперсионный анализ объединяет все эти сравнения. Конечно, можно начать с проверки F -критерия, а затем использовать отдельные дополнительные критерии для интересующих нас сравнений. Однако существует риск, что главный эффект F не будет обладать значимостью, и в этом случае у нас не будет реальных оснований для дополнительных проверок. Кроме того, критерии, используемые после получения значимого результата при дисперсионном анализе, обладают меньшей мощностью, что означает риск упустить из вида действительно важный эффект.

Альтернативой ДА в таких случаях служат плановые сравнения, когда мы заранее определяем, какие из средних значений будем сравнивать, и производим толь-

ко эти сравнения. В нашем гипотетическом исследовании обучения, к примеру, мы могли бы провести сравнение экспериментальных воздействий, оставив без внимания другие сравнения. Разъяснение методов проведения такого избирательного сравнения не входит в намерения автора; описание их можно найти в большинстве учебников по статистике (например, Hays, 1981). Разумеется, такой подход влечет потерю некоторой информации. Однако если мы точно знаем, что нас интересует, потеря будет минимальной. А поскольку критерии, используемые в предварительно спланированных проверках, обладают большей мощностью, чем критерии *post-hoc*, наши шансы получить ясные вопросы на интересующие нас вопросы выше.

Важно подчеркнуть, что плановые сравнения действительно подразумевают планирование и избирательность; то есть мы не можем сравнивать что угодно. Каким должно быть количество сравниваемых пар — на этот счет специалисты в области статистики не имеют единого мнения. Одни из них рекомендуют ограничивать плановые сравнения статистически независимыми, или «ортогональными», парами. Количество таких независимых пар на одну меньше количества средних значений; в нашем исследовании обучения оно будет составлять 5 независимых сопоставлений (в упоминавшейся выше работе объясняется, как определить независимость сопоставлений (Hays, 1981)). Другие специалисты являются приверженцами несколько более либерального подхода, говоря о том, что интересные с теоретической точки зрения пары можно проверить даже при отсутствии полной их независимости. В работе Кеппел (Keppel, 1991) можно найти полезную информацию о разных точках зрения, а также ряд методов корректировки уровня вероятности в случаях, когда производятся множественные или частично пересекающиеся сравнения.

Величина эффекта

Цель плановых сравнений состоит в выявлении эффектов, которые могут упустить такие глобальные проверочные процедуры, как ДА. Процедура измерения величины эффекта снимает некоторые из ограничений ДА. В этом случае мы предполагаем, что при анализе был обнаружен значимый эффект; тогда встает вопрос, насколько он велик. Насколько сильна связь между зависимыми и независимыми переменными?

Чтобы разобраться в этом вопросе, нужно вспомнить то, о чем мы говорили ранее, обсуждая термин «статистическая значимость». Установление факта статистической значимости свидетельствует о наличии и некой неслучайной связи между переменными. Факт наличия значимости ничего не говорит о силе этой связи. О масштабах эффекта можно, конечно, догадываться по средним значениям; большая разница между средними, очевидно, отражает более значимый эффект, чем меньшая разница. Но есть ли более точная мера величины эффекта?

Ответ состоит в том, что сегодня для расчета величины эффекта существует ряд методов. Основные работы, в которых дается описание этих методов, следующие: Коэн (Cohen, 1977), Розенталь (Rosenthal, 1994b) и Тацоука (Tatsouka, 1993). Здесь я опишу простейшую из процедур, разработанную Коэном (Cohen, 1977). Согласно этому подходу, величина эффекта, или d , определяется как разница между двумя средними, разделенная на стандартное отклонение в сравниваемых группах.

Таким образом, учитывается средняя разница, которая оценивается с точки зрения изменчивости показателей. Чем меньше изменчивость, тем существеннее средняя разница.

Сейтц (Seitz, 1984) приводит пример. Стандартное отклонение в большинстве тестов IQ равно 15. Средняя разница между группами, равная 12, (каковой она, к примеру, указывается в исследовательских отчетах для только что поступивших в колледж и докторов философии) означала бы величину эффекта 0,8, если бы средняя разница составляла 7 пунктов, величина эффекта была бы равной 0,2. Для интерпретации этих величин можно графически изобразить распределение показателей двух популяций и область их пересечения. На рис. 7.5 изображены кривые, соответствующие трем описанным выше ситуациям. Заметьте, что с возрастанием величины эффекта сокращается область совпадения. Как отмечает Сейтц, средняя пара кривых особенно информативна. Средняя разница 7 пунктов — или, если рассматривать ситуацию в общем виде, разница, составляющая половину стандартного отклонения, — может показаться не слишком большой. Однако эта разница означает, что 70 % одной популяции имеет показатели выше среднего другой популяции.

Меры величины эффекта дают полезную информацию, которую нельзя извлечь непосредственно из значений логических критериев. Однако до сих пор в исследовательских отчетах в области психологии развития — а в действительности и в психологии в целом — редко можно встретить указания на величину эффекта (Cohen, 1994). При обзоре любого журнала по психологии развития можно обнаружить не один десяток F и t , но лишь несколько скромных попыток рассчитать формальные показатели величины разнообразных эффектов.

Мультивариантный дисперсионный анализ

Разница между одновариантными и мультивариантными статистическими процедурами довольно расплывчата; разные специалисты определяют ее по-разному. Однако наиболее распространенным критерием служит количество зависимых переменных. Если анализ проводится при наличии одной зависимой переменной, его можно назвать одновариантным. Процедура использования t -критерия и дисперсионный анализ — одновариантные статистические процедуры. Если в анализе задействуется более одной зависимой переменной, его можно назвать мультивариантным. Существует ряд мультивариантных статистических процедур, к которым относится и мультивариантный дисперсионный анализ, или МДА.

Использование мультивариантных статистических процедур — явление, получившее широкое распространение в психологических исследованиях совсем недавно. Как отмечает Сейтц (Seitz, 1980), статистические основы для этих процедур были выработаны уже давно; проблемы начинались при их практическом использовании. Расчеты мультивариантных статистических показателей зачастую чрезмерно сложны и громоздки, и поэтому до появления компьютеров и соответствующих компьютерных программ использование мультивариантных статистических процедур было крайне затруднено. Наличие компьютера и навыки работы на нем могут оказаться весьма ценными для проведения любой формы статистического анализа.

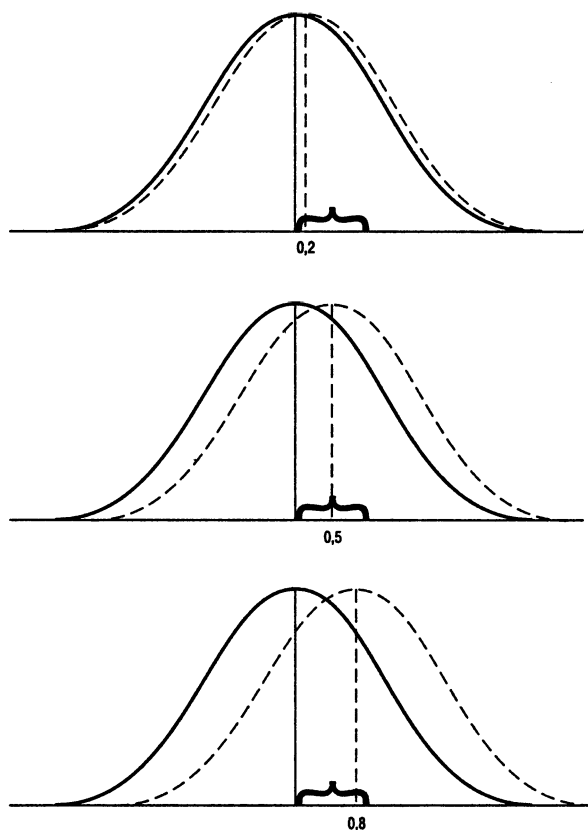


Рис. 7.5. Разница между популяциями, соответствующая разным значениям величины эффекта

Расчеты — не единственная сложность, связанная с мультивариантными статистическими процедурами. В определенном смысле, расчеты — самое простое, поскольку эту работу выполняет компьютер. Настоящая проблема — определить, когда необходим мультивариантный анализ и как интерпретировать его результаты. Для ответа на эти вопросы написано множество книг и прочитано множество теоретических курсов, в том числе весьма подробных и содержательных (см., например: Hair, Anderson, Tatham & Black, 1992; Morrison, 1990; Nesselroade & Cattell, 1988). Здесь я затрону лишь несколько моментов, о которых ведется речь в более сжатой работе (Applebaum & McCall, 1983), ориентированной на специалистов в области психологии развития.

Предположим, проведено исследование с двумя или более зависимыми переменными. Как узнать, нужно ли проводить отдельный ДА для каждой из переменных (этот вариант всегда необходимо учитывать) или объединить их все в МДА? Эпплбаум и Маккол говорят о двух основных преимуществах МДА перед ДА. Одно из них сходно с преимуществом F -критерия перед отдельными t -критериями. Чем больше проверок с помощью дисперсионного анализа мы проводим, тем выше вероятность, что некоторые эффекты достигнут уровня статистической зна-

чимости чисто случайно. При проведении МДА уровень вероятности при любых сравнениях, напротив, сохраняется постоянным. Второе преимущество имеет отношение к возможности при проведении МДА использовать информацию о связи между зависимыми переменными, информацию, которая не учитывается при ДА, когда анализ каждой переменной производится отдельно. Одновременный учет всех переменных, среди прочего, означает, что при помощи МДА иногда можно выявить значимые эффекты, которые упускает ДА.

Как следует из вышесказанного, МДА уместно использовать только тогда, когда есть основания предположить, что между зависимыми переменными существует некая интересная связь. Как пишут Эпплбаум и Маккол (Applebaum & McCall, 1983): «Вы не включаете в МДА все имеющиеся переменные, лишь бы увидеть, что из этого получится. Зависимые переменные нужно осмысленно подбирать. Они должны образовывать логическую совокупность, которую можно интерпретировать именно как совокупность» (р. 435). Затем Эпплбаум и Маккол приводят в качестве примера исследование габитуации у младенцев, в котором зависимые переменные представляют собой следующее: длительность фиксации в первой пробе, пробу с наиболее длительной фиксацией, количество проб от самой длительной к самой непродолжительной фиксации и средняя длительность фиксации в одной пробе. Эти переменные образуют совокупность связанных между собой (но не дублирующих друг друга) показателей габитуации у младенца, для анализа которой больше подходит МДА, а не ДА. В целом, для использования МДА больше всего подходит ситуация, когда исследование включает ряд несколько различающихся между собой характеристик одного и того же конструкта (как в примере с габитуацией).

Как бы ни было кратко это изложение метода МДА, из него можно вынести главную мысль. Наличие мощной статистической процедуры — и компьютерной технологии для ее осуществления — не означает, что эту процедуру можно бездумно использовать для любого типа данных. Сутью грамотной статистической работы является знание того, какой статистический прием применим к разного рода данным и планирование исследования с целью достижения оптимального соответствия статистических процедур и данных.

Множественный регрессионный анализ

Как и МДА, множественный регрессионный анализ — сложная статистическая процедура — сложная как для проведения, так и для описания. Как и в случае с МДА, на двух-трех страницах излагать суть процедуры множественного регрессионного анализа бессмысленно. Однако регрессионный анализ стал использоваться столь широко, что краткое рассмотрение его принципов представляется оправданным.

Говоря словами Керлингер (Kerlinger, 1986), множественный регрессионный анализ — «это метод изучения результатов влияния и силы влияния более чем одной независимой переменной на одну зависимую переменную с использованием принципов корреляционного и регрессионного анализа» (р. 527). Если выразиться несколько иначе, множественный регрессионный анализ дает нам информацию о том, как связаны две или более независимые переменные, или «предикторы», с

одной зависимой переменной, или «критерием». Мы, к примеру, можем провести исследование, в котором будем изучать успешность выполнения в лабораторных условиях некоего задания как функцию от IQ , социально-экономического статуса, характера инструкций к заданию и времени, отведенного на его выполнение. Применение множественного регрессионного анализа к нашим данным позволит оценить вклад каждого из этих четырех предикторов (как по отдельности, так и в сочетании) в дисперсию критериальной переменной.

Между регрессионным и корреляционным анализом, описанным ранее в этой главе, существует тесная связь. В обоих случаях мы используем значение одной переменной у некоего испытуемого для предсказания значения другой переменной. Как следует из факта включения таких параметров, как IQ и социально-экономический статус, еще одно сходство состоит в возможности использования в регрессионном анализе переменных, которые, с точки зрения исследовательского плана, являются корреляционными, в том смысле, что они просто измеряются, а не подвергаются экспериментальным манипуляциям.

Много общего у регрессионного анализа и с дисперсионным анализом. Эти процедуры имеют одни и те же цели: определить влияние, совокупности независимых переменных на некую зависимую переменную, как по отдельности, так и во взаимодействии. Регрессионный анализ в предыдущем примере можно было бы заменить дисперсионным анализом четырех переменных, результатом которого было бы значение четырех F , указывающих на главный эффект, а также ряд значений F для взаимодействий. В действительности это правило носит общий характер: большинство задач, которые решает множественный регрессионный анализ, можно решить с помощью дисперсионного анализа. Однако, как отмечают многие авторы, из двух процедур множественный регрессионный анализ обладает более широкими возможностями, поскольку ДА — это лишь частный случай регрессионного анализа. Поэтому любую задачу, которую может решить многофакторный ДА, может решить и множественный регрессионный анализ.

Если учесть широкие возможности множественного регрессионного анализа, можно задаться вопросом, какие факторы могут заставить нас предпочесть его ДА при анализе некой совокупности данных? Отмечу вначале, что зачастую это дело вкуса, поскольку во многих случаях оба подхода равно эффективны и информативны. Однако в некоторых отношениях множественный регрессионный анализ обладает преимуществом. Здесь я приведу два аргумента в пользу его использования.

1. Множественный регрессионный анализ особенно удобен в тех случаях, когда независимая переменная является непрерывной величиной, то есть включает широкое множество значений, а не просто несколько дискретных уровней. Пример непрерывной переменной — IQ . Включив в наше исследование IQ , мы получили бы разброс значений около 60–70 пунктов. При использовании ДА учесть всю эту дисперсию невозможно; самое большее, что мы могли бы сделать, это произвести приблизительную классификацию на «высокий», «средний» и «низкий» интеллект. При использовании же множественного регрессионного анализа нет потери информации, поскольку переменные рассматриваются на континууме и анализируются все фактические показатели.

2. Множественный регрессионный анализ особенно подходит для изучения вопроса, о котором уже говорилось в этом разделе: определения силы влияния независимой переменной на зависимую. Основным статистическим показателем во множественном регрессионном анализе является R^2 : доля дисперсии зависимой переменной, объясняемая совокупностью изучаемых независимых переменных. То есть R^2 указывает на то, как «действуют» наши предикторы — насколько точно мы можем предсказать изменчивость критерия при помощи выбранной совокупности независимых переменных. Хотя сделать это и нелегко, при помощи регрессионного анализа все же можно оценить вклад каждого из предикторов в дисперсию критерия, как количественное выражение этого вклада, так и тип связи предиктора с критерием.

Среди учебников, посвященных множественному регрессионному анализу, можно назвать такие работы, как Cohen & Cohen, 1983; Draper & Smith, 1981; Pedhazur, 1982. В книге Керлингера (Kerlinger, 1986), на которую мы уже ссылались, можно найти весьма доступное введение в данную проблематику; кроме того, это великолепная работа по методологии, с которой стоит ознакомиться любому, кого интересует тема психологического исследования.

Общие принципы

Завершим эту главу некоторыми общими рекомендациями, касающимися статистической части исследования.

Первая рекомендация уже была дана в главе 1: все необходимо планировать. Из всего сказанного в этой главе должно быть ясно, что есть такие аспекты анализа данных, которые нельзя полностью предусмотреть. Исследователь, к примеру, может не знать, каким будет распределение данных, пока не осуществит их сбор. Могут также появляться неожиданные, но интересные результаты, требующие статистической проверки, которую нельзя предусмотреть с самого начала. Однако настолько, насколько это возможно, количество таких статистических операций по уже свершившемуся факту следует свести к минимуму. С самого начала исследователь должен знать основные вопросы, которые будут анализироваться, и конкретные статистические критерии, которые будут использованы при анализе. Планирование помогает избежать несвоевременного разбора данных, с проведением множества проверок в поисках того, что могло бы оказаться значимым. Кроме того, планирование позволяет удостовериться в том, что для анализируемых данных и вопросов есть подходящие статистические процедуры, и выбрать наиболее мощные и информативные критерии.

Вторая рекомендация — будьте аккуратны. Выводы, которые можно сделать на основе результатов исследования, полностью зависят от точности статистических расчетов. Ошибка при вычислении может привести к ошибке в выводах, иногда принципиальной. Кроме того, вычислительные ошибки обнаружить труднее, чем другие; в исследовательском отчете можно обнаружить методологические просчеты, но результаты математических действий читатель, как правило, принимает на веру. Поэтому крайне важно, чтобы каждая операция проходила тщательную проверку. В идеальном случае, все расчеты должны произвести, по меньшей мере, два

независимых специалиста. Полезно также сравнить результаты двух или более методов подсчета (например, на калькуляторе и на компьютере).

От исследователя требуется не просто обычная аккуратность; иногда возникают ситуации, когда статистические результаты столь необычны, что их нельзя не перепроверить. В частности, подозрения должны вызывать любые случаи несоответствия дескриптивных показателей и логических выводов о значимости. Дескриптивные и логические показатели основаны на одних и тех же данных, и поэтому они всегда должны определенным образом согласовываться. Установление статистической значимости ничтожной разницы или отсутствие значимости довольно существенной разницы должно быть стимулом к перепроверке расчетов.

Есть и еще одно замечание, которое можно сделать о проверке. Исследователь должен избегать дифференцированных проверок, когда результаты, противоречащие ожиданиям, анализируются со всей тщательностью, а результаты, подтверждающие ожидания, не проверяются. Как бы ни было естественным такое поведение, оно позволяет вкрасься эффектам необъективности экспериментатора: результатом исправления только негативных ошибок может быть псевдодоказательство гипотез исследователя. Разумеется, лучше всего проверять и корректировать все.

Последняя рекомендация — обращайтесь за советом. Сегодня доступно столько источников — учебники, компьютерные программы, специалисты в области статистики, — что новичку нужно быть совершенно безрассудным человеком, чтобы в одиночку пуститься в плавание по коварным водам статистики. К эксперту можно обратиться на любом этапе процесса исследования, начиная с первичного выбора системы измерения и плана и кончая подготовкой к публикации исследовательского отчета.

Возможно, имеет смысл несколько более подробно рассмотреть письменные источники. Я уже упоминал лучшие учебники по всему курсу статистики, а также по отдельным его темам. Некоторые статистические вопросы и соответствующие процедуры имеют особое значение в исследованиях в области психологии развития. В этой главе мы не рассматривали статистические процедуры, наиболее часто используемые в психологии развития; этому есть два объяснения: большинство вопросов статистики, с которыми имеет дело исследователь в области психологии развития, в действительности ничем не отличаются от вопросов, с которыми сталкиваются все исследователи-психологи; а те вопросы, которые все же специфичны для психологии развития, настолько сложны, что углубляться в них здесь не имеет смысла. Заметьте, однако, что есть полезные (хотя подчас весьма трудные) работы, посвященные статистическим процедурам, имеющим особое значение именно в психологии развития (см. например: Achenbach, 1978; Applebaum & MacCall, 1983; Collins & Horn, 1991; Nunnally, 1982).

Резюме

Психологи используют статистические показатели, преследуя две связанные между собой цели. Дескриптивные статистические показатели дают краткую характеристику данным; они представляют собой своего рода перечное описание того, что было обнаружено в исследовании. К ним относят меры центральной тенденции,

главной из которых является среднее, и меры изменчивости, главными из которых являются дисперсия и стандартное отклонение.

С помощью статистических показателей, выводимых логическим путем, производится не просто описание данных, а определение статистической значимости. Суть вопроса состоит в том, значимо ли отклоняются полученные результаты от того, что могло произойти случайно, при этом случайность является основой нуль-гипотезы об отсутствии различий между сравниваемыми группами. В качестве примера логических критериев рассматривается t -критерий, который можно использовать для сравнения средних значений в двух группах. Как и у большинства логических критериев, значимость t -критерия зависит от трех факторов: степени различия групп, уровня изменчивости значений в каждой группе и объема выборки. Подчеркивается, что выводы, основанные на использовании логических критериев, всегда носят вероятностный характер; всегда существует возможность допустить ошибку: ошибку первого рода, при которой происходит ошибочное отвержение истинной нуль-гипотезы, и ошибку второго рода, при которой не отвергается ложная нуль-гипотеза. Подчеркивается также, что при проверке статистической значимости исключается возможность объяснения результатов случайной дисперсией, однако сама статистическая значимость не гарантирует ни валидности исследования, ни научной ценности результатов.

Затем рассматриваются виды статистического анализа, которые подходят для разного типа данных, то есть вопрос, как выбрать логический критерий. При выборе критерия следует учитывать три фактора: первый — это уровень измерения — производится измерение по номинальной, порядковой, интервальной шкале или по шкале интервалов. Второй фактор — распределение полученных данных — в частности, распределены ли данные по закону нормального распределения. Так называемые параметрические критерии, например t -критерий и F -критерий, зависят от определенных предположений, касающихся распределения данных. Часто они обладают большей мощностью, чем непараметрические критерии, лучше выявляя истинные эффекты. В то же время возможность использования параметрических критериев ограничена рядом условий. Третий фактор — план исследования: сколько включено независимых переменных, сколько уровней имеет каждая переменная, производится внутри- или межсубъектное сравнение.

Помимо выявления межгрупповых различий существует еще одна область использования статистических процедур — установление меры связи между переменными. Корреляционные статистические показатели указывают на степень линейной связи между переменными. Статистическая значимость коэффициента корреляции зависит от его величины и объема выборки. Сила связи зависит от величины коэффициента корреляции: чем ближе он к единице, тем больше возможностей предсказать один показатель, зная другой. На коэффициент корреляции влияет диапазон изменения показателей в двух группах; чтобы результаты можно было генерализировать, этот диапазон должен соответствовать диапазону изменения показателей в популяции.

Как бы ни был гибок дисперсионный анализ, есть ситуации, когда уместнее использовать другую статистическую процедуру. В исследованиях с четкими гипотезами предпочтительнее использовать плановые сравнения, а не обобщенный ДА.

Предварительно спланированные процедуры более эффективны, чем единичные сравнения. Наряду со значимостью исследователя часто интересует сила влияния независимой переменной. Существуют разнообразные меры величины эффекта, простейшая из них — d — разница между двумя средними, разделенная на стандартное отклонение. В исследованиях с двумя и более зависимым переменными возможной альтернативой ДА является мультивариантный дисперсионный анализ, или МДА. МДА используется в тех случаях, когда зависимые переменные образуют совокупность связанных между собой элементов, при этом МДА дает возможность осуществить более точную проверку эффектов. Наконец, альтернатива ДА — множественный регрессионный анализ, система методов статистического анализа, одним из которых, собственно, и является ДА. Хотя две эти процедуры зачастую взаимозаменяемы, есть ситуации, когда множественный регрессионный анализ предпочтительнее ДА, в частности, когда независимые переменные — непрерывные, а не дискретные величины.

Главу завершает изложение нескольких общих замечаний — три рекомендации: составлять предварительный план, производить расчеты аккуратно и при необходимости обращаться к специалистам.

Упражнения

Исследователь А сравнивает 10 экспериментальных и 10 контрольных испытуемых и выявляет значимое различие между группами на уровне $p = 0,04$. Исследователь В сравнивает 100 экспериментальных и 100 контрольных испытуемых и выявляет значимое различие между группами на уровне $p = 0,04$. О каком из этих различий вы бы смогли с большей уверенностью сказать, что это не результат случайности? Какие из результатов отражают более серьезное различие между условиями проведения исследования? Обоснуйте свои выводы.

Как вам стало известно из этой главы и как вы увидите в дальнейшем, когда будете читать раздел, посвященный результатам, обозначение статистических показателей несет определенный смысл. Соответственно, анализ кратких обозначений статистических терминов помогает проверить, насколько верно вы понимаете смысл этих терминов и понятий. Объясните как можно более точно, что, по вашему мнению, обозначают понятия, стоящие за представленными ниже обозначениями:

p r
 t N
 F SD .

Глава 8

Этика

В главе 7 подчеркивалось, что вопросы статистической обработки не стоит откладывать до заключительной фазы исследования. То же замечание справедливо, и даже более существенно, в отношении вопросов этики. Любое исследование должно отвечать двум требованиям: научности и этической безупречности. Если по поводу хотя бы одного из этих критериев возникают серьезные сомнения, исследование проводить нельзя.

Тридцать-сорок лет назад решение вопросов этики оставалось преимущественно на совести исследователя. Сегодня все иначе. В последние годы происходила разработка все новых и новых норм и принципов проведения исследований с людьми. Кроме того, сегодня права испытуемого имеют множество степеней защиты. При работе с детьми нормы и правила становятся даже еще строже.

Эту главу я начну с обзора норм, в соответствии с которыми должен действовать исследователь. Во вводном разделе в общих чертах излагается порядок проверки исследовательского проекта на соответствие этическим принципам. Затем мы более подробно рассмотрим основные права, которыми обладает каждый, кто принимает участие в исследовании в качестве испытуемого. Особое внимание уделяется трем правам: праву на информированное согласие, на гарантии безопасного участия в исследовании и праву на конфиденциальность информации, полученной в ходе исследования. Детально упомянутые вопросы рассмотрены у Сибера (Sieber, 1992), Стэнли и Сибера (Stanley & Sieber, 1991), а также у Томпсона (Thompson, 1990). В Фишер и Трайтон (Fisher & Tryton, 1990) делают более глубокий анализ этических вопросов, актуальных для прикладных исследований в области психологии развития.

Прежде чем обратиться к рассмотрению норм и процедур, сделаем одно замечание. Следование установленным правилам — необходимое условие этичности исследовательского проекта. Однако этого недостаточно. Все нормы неизбежно носят несколько отвлеченный характер, и исследователь должен постоянно примерять их к своему конкретному случаю. Этический кодекс помогает при принятии решения, но не заменяет его. Кроме того, сколько бы инстанций ни одобрило проект, конечная ответственность за этичность исследования всегда лежит на самом исследователе.

Нормы и процедуры

Предположим, ученый задумал исследование, которое он намеревается провести на испытуемых-школьниках младших классов. Он вполне доволен своим проектом с научной точки зрения, однако хочет удостовериться в том, что все запланированные процедуры не противоречат этическим принципам. В этом разделе мы рассмотрим два вопроса: Откуда исследователь может почерпнуть информацию об этических нормах? И что он должен сделать, прежде чем начать сбор данных?

О нормах этического поведения при проведении исследования можно узнать из разнообразных источников. (Более подробно об историческом развитии этих стандартов см.: Cooke, 1982 и Rheingold, 1982). Общие международные стандарты сформулированы в Нюрнбергском кодексе (Trials of War Criminals..., 1949) и в Хельсинкской декларации (World Medical Association, 1964). Правительство США также регулирует проведение исследований с испытуемыми-людьми, преимущественно посредством правил, которыми руководствуются разнообразные агентства, распределяющие правительственные гранты на исследования. Американская психологическая ассоциация, к примеру, опубликовала «Этические принципы психологов и кодекс поведения» (Ethical Principles of Psychologists and Code of Conduct, 1992), а Канадская психологическая ассоциация издала Канадский кодекс этики для психологов (Canadian Code of Ethics for Psychologists, 1991). Кроме того, Общество исследований в области детского развития (Society for Research in Child Development), главная междисциплинарная организация исследователей в области детского развития, опубликовала нормы проведения исследований с детьми (Committee for Ethical Conduct, 1990).

Упомянутые выше издания, разумеется, не являются абсолютно независимыми друг от друга работами. Декларируемые в них принципы сходны и нередко идентичны, а авторы нередко заимствуют друг у друга отдельные параграфы. В связи с тем, что этический кодекс, опубликованный Обществом исследований в области детского развития, представляет для специалистов в области психологии развития наибольший интерес, его положения представлены во Вставке 8-1.

Вернемся теперь к нашему гипотетическому исследователю. Допустим, что он знаком — и это, собственно, его обязанность — с принципами, сформулированными в упомянутых выше документах. На местном уровне выполнение этих принципов контролирует организация, называемая **институциональным наблюдательным советом** (*institutional review board*). Через этот совет и будет действовать исследователь.

Существование наблюдательных советов отражает один из основных принципов исследовательской этики: принцип **независимой оценки**. Идея, положенная в основу независимой оценки, заключается в том, что лицо, планирующее исследование, имеет слишком большую заинтересованность в исследовании, чтобы быть беспристрастным при определении его этичности. Необходим анализ исследования независимыми компетентными лицами, которые не имеют никакого отношения к исследованию. С этой целью во всех университетах, где проводятся исследования с людьми, создаются институциональные наблюдательные советы, в обязанности которых входит контроль этичности выдвигаемых исследовательских проектов.

Члены этих советов являются представителями разных научных дисциплин. Одно из условий, сформулированных в правительственном постановлении о создании подобных советов, звучит следующим образом: «Члены наблюдательных советов не могут быть непосредственно заинтересованы в оцениваемом исследовании, совет не может целиком состоять из сотрудников исследовательского учреждения, лиц, его представляющих, или членов одной профессиональной группы» (Cooke, 1982, p. 167).

Каждый институциональный наблюдательный совет публикует свой собственный перечень норм, близких к нормам, изложенным в правительственных и профессиональных документах, упоминавшихся выше. Кроме того, каждый совет устанавливает стандартную форму заявления для рассмотрения исследовательского проекта. Как правило, этот бланк начинается с вопросов о целях исследования и этапах его проведения. Если совет сочтет необходимым, то может запросить копии экспериментальных протоколов, опросников и т. д. Помимо этой информации общего характера в бланк может быть включен ряд специфических вопросов, касающихся этичности исследования. Эти вопросы (о которых мы вскоре поговорим более подробно) состоят в следующем. Представляет ли исследование какую-либо угрозу для испытуемого? Какую пользу получит испытуемый, приняв участие в исследовании? Как испытуемые будут привлекаться к участию в исследовании, и будут ли использованы какие-либо формы поощрения за сотрудничество? Что конкретно будет сказано испытуемым об исследовании до получения их согласия на участие? В тех случаях, где это имеет смысл, узнают ли испытуемые в конце исследования об обманных приемах, к которым прибегал исследователь, и об истинных целях исследования? Каким образом будет оберегаться конфиденциальность результатов?

Одобрение наблюдательного совета дает исследователю право претендовать на грант или, если исследования не финансируются, начать поиски испытуемых и проводить исследование. Если говорить о нашем гипотетическом исследователе младших школьников, ему, скорее всего, придется работать через систему государственных школ. Конкретный порядок действий зависит от того, где осуществляется исследование, хотя есть некоторые общие моменты. В некоторых школьных округах есть главное управление, через которое проходят все исследовательские заявки, причем заявки передаются в школу только после прохождения этой первой инстанции. Заявку, попавшую в школу, должен одобрить директор. В одних школах директор принимает решение самостоятельно, в других — директор советуется с учителями, классы которых будут принимать участие в исследовании, и дает свое одобрение только по договоренности с учителями. Достигнув уровня класса, заявка передается родителям, как правило, в форме письменного обращения, в котором описывается суть исследования и испрашивается разрешение на участие ребенка в исследовании. Обследовать можно только тех детей, родители которых вернули бланки с обращением подписанными. Наконец, ребенок должен получить представление о том, в чем будет состоять исследование, и выбрать, участвовать в исследовании или нет.

Как явствует из приведенного выше описания процедуры, между исследовательским замыслом и реальным обследованием детей стоит ряд охранительных

Этические стандарты исследований детей (Общество исследований в области детского развития, 1990)

Доклад Комитета по этическому поведению при проведении исследований в области детского развития.
Принципы

Принцип 1. Нанесение ущерба испытуемому. Исследователь должен воздерживаться от любых действий, которые могут нанести ребенку либо физический, либо психологический вред. Исследователь также обязан всегда, когда это возможно, использовать наименее стрессогенные процедуры. В некоторых случаях психологический вред выявить трудно: тем не менее, обязанностью исследователя остается выявление негативных последствий, создание условий для их смягчения или полное их устранение. Если у исследователя имеются сомнения относительно возможности пагубного воздействия исследовательских процедур, необходимо обратиться за консультацией к другим специалистам. Если причинение вреда представляется неизбежным, исследователь обязан найти другие способы получения интересующей его информации или отказаться от исследования. Тем не менее бывают ситуации, когда подвергание ребенка стрессогенным процедурам необходимо в диагностических или терапевтических целях. В таких случаях требуется взвешенная оценка Институционального наблюдательного совета.

Принцип 2. Информированное согласие. До получения согласия или разрешения ребенка на участие в исследовании необходимо проинформировать его обо всех особенностях исследования, которые могут повлиять на его решение об участии, а также ответить на все вопросы ребенка в доступной для него форме. Исследователь должен уважать право ребенка выбирать, принимать ему участие в исследовании или нет, предоставляя возможность согласиться или не согласиться на участие, а также в любой момент выйти из исследования. Под разрешением (*assent*) подразумевается та или иная форма положительной оценки ребенком перспективы своего участия в исследовании, не обязательно при полном понимании смысла исследования, необходимом для того, чтобы дать информированное согласие (*informed consent*). Исследователи, работающие с младенцами, должны особенно подробно обсудить с родителями суть исследования и особенно внимательно относиться к любым проявлениям младенцем ощущения дискомфорта.

Несмотря на первостепенное значение получения информированного согласия, иногда согласие испытуемого или любая форма контакта с ним делает проведение исследования невозможным. Характерный пример — полевые исследования методом невмешательства. Такие исследования можно считать этическими, если они проводятся в общественных местах, с полной защитой анонимности и отсутствием каких-либо известных негативных последствий для испытуемых. Однако решение о том, этично ли проведение такого исследования в конкретных обстоятельствах, должно выноситься с учетом мнения Институционального наблюдательного совета.

Принцип 3. Согласие родителей. Исследователь должен получить информированное согласие родителей ребенка, законных опекунов или тех, кто выполняет их функции (например, учителя, директора учебного заведения и т. д.), предпочтительно в письменной форме. Информированное согласие предполагает, что родители или другие ответственные лица осведомляются обо всех особенностях исследования, которые могут повлиять на их решение позволить или не позволить ребенку участвовать в исследовании. Исследователь должен также предоставить информацию о своей специальности и месте работы. Необходимо не просто уважать право ответственных за ребенка лиц отказаться в согласии, но и проинформировать их о том, что их отказ не повлечет никаких неприятностей ни для них, ни для ребенка.

Принцип 4. Дополнительное согласие. Исследователь должен получить информированное согласие любых лиц, например школьных учителей, взаимодействие которых с ребенком является предметом исследования. Как и в случае с ребенком, родителями или опекунами, информированное согласие предполагает, что лица, взаимодействующие с ребенком в ходе исследования, осведомляются обо всех особенностях исследования, которые могут повлиять на их решение принимать в нем участие или нет. Следует ответить на все вопросы этих лиц и уважать их право участвовать или не участвовать в исследовании, а также право в любой момент отказаться от участия в исследовании.

Принцип 5. Поощрение. Поощрение за участие в исследовании должно быть разумным и не превышать по своей значительности тех поощрений, которые ребенок получает в обычной жизни. Какого бы

рода поощрение не предлагалось, необходимо всегда помнить о том, что, чем сильнее исследование может отразиться на ребенке, тем большую ответственность за благополучие и свободное волеизъявление ребенка несет исследователь.

Принцип 6. Обман. Хотя полное раскрытие информации при получении согласия является этическим идеалом, для проведения исследования может потребоваться сокрытие части информации или обман. Считая необходимым частичное сокрытие информации или обман, исследователь должен убедить своих коллег в правомерности своего решения. При использовании сокрытия информации или обмана в ситуациях, когда есть основания полагать, что это негативно отразится на испытуемых, после проведения исследования необходимо принять соответствующие меры по информированию испытуемых о мотивах обмана. Если исследование построено на сокрытии информации, необходимо использовать такие обманные приемы, которые не имеют известных негативных последствий для ребенка или его семьи.

Принцип 7. Анонимность. Для получения доступа к документам, исследователь должен обратиться за разрешением к лицу, ответственному за эти документы. Необходимо сохранить анонимность этой информации и использовать только ту часть информации, на использование которой получено разрешение. Исследователь обязан удостовериться в том, что указанное выше ответственное лицо действительно пользуется доверием испытуемого и, давая разрешение на использование информации, берет на себя определенную ответственность.

Принцип 8. Взаимные обязательства. С самого начала любого исследования между исследователем и родителями, опекунами или теми, кто выполняет их функции, а также ребенком должно существовать ясное соглашение, определяющее обязанности каждой стороны. Исследователь должен выполнить все данные обещания и обязательства, предусмотренные этим соглашением.

Принцип 9. Опасность. Если в ходе исследования исследователь узнает о чем-либо, что угрожает благополучию ребенка, он должен обсудить этот вопрос с родителями или опекунами и специалистами в данной области с тем, чтобы оказать ребенку необходимую помощь.

Принцип 10. Непредвиденные последствия. Если исследовательские процедуры неожиданно привели к нежелательным для испытуемого последствиям, исследователь должен немедленно принять соответствующие меры по исправлению этих последствий и пересмотреть порядок этих процедур в том случае, если он намеревается использовать их в дальнейшем.

Принцип 11. Конфиденциальность. Исследователь должен сохранять в тайне информацию, полученную об испытуемых. В письменных и в устных отчетах об исследовании, а также в неформальных беседах со студентами и коллегами должны скрываться любые сведения, позволяющие установить личность испытуемого. Если существует вероятность того, что к подобного рода сведениям получат доступ посторонние лица, эта вероятность, а также меры по сохранению конфиденциальности должны быть разъяснены испытуемым в ходе получения информированного согласия.

Принцип 12. Информирование испытуемых. Сразу после сбора данных исследователь должен прояснить для испытуемых все вопросы, которые могли у них возникнуть в ходе исследования. Исследователь также должен в доступной для испытуемых форме изложить основные выводы из исследования. Там, где научные или общечеловеческие принципы оправдывают сокрытие части информации, необходимо предпринять все возможное, чтобы это не имело для испытуемых никаких негативных последствий.

Принцип 13. Сообщение результатов. Поскольку слова исследователя могут иметь для родителей и детей последствия, о которых исследователь не догадывается, сообщая результаты, давая оценки и рекомендации, необходимо соблюдать известную осторожность.

Принцип 14. Ответственность ученого за последствия научных открытий. Исследователь всегда должен помнить о социальном, политическом и общечеловеческом значении своего исследования и с особой осторожностью представлять результаты своей работы. Этот принцип, однако, не лишает исследователя права работать в любой научной области или права соблюдать правила предоставления научного отчета.

Источник: «SRCED Ethical Standards for Research with Children», Society for Research in Child Development, 1990, Newsletter, p. 5–6.

барьеров. Основное назначение этих барьеров, конечно же, гарантирование прав ребенка. С точки зрения исследователя, эти меры защиты имеют дополнительный смысл. Во-первых, важно не только как следует продумать меры по обеспечению этичности при планировании исследования, но и убедительно доказать этическую безупречность своего проекта другим людям. Исследование может осуществиться только в том случае, если в его ценности и этичности будет уверено множество людей. Нужно предусмотреть время для получения одобрения различных инстанций. В зависимости от того, насколько удачным будет сотрудничество с теми или иными учреждениями, между этапом планирования исследования и началом сбора данных может пройти несколько недель, а то и месяцев.

Следует еще раз подчеркнуть один момент. В этом разделе особое внимание было уделено мероприятиям, гарантирующим соблюдение прав ребенка, принимающего участие в исследовании. Однако тот факт, что столько инстанций имеет возможность защитить ребенка, не снимает личной ответственности исследователя. Как уже было сказано, в конечном счете, ответственность за соблюдение этических норм всегда лежит на том, кто проводит исследование.

Права испытуемых

Информированное согласие

Участие в исследовании должно быть добровольным. Нельзя проводить обследование без ведома объекта обследования; нельзя также принуждать к участию в исследовании. Эти принципы объединяет в себе положение об **информированном согласии**: испытуемые должны заранее знать, в чем будет состоять исследование, и дать свое недвусмысленное согласие на участие¹.

Принцип информированного согласия достаточно легко декларировать, но осуществление его на практике связано с массой трудностей. Именно об этих трудностях мы и поговорим. Начнем с более подробного рассмотрения компонентов этого понятия: «информированное» и «согласие». Затем перейдем к вопросам, непосредственно связанным с обследованием детей.

Определение «информированное» подразумевает, что простого согласия на участие недостаточно; прежде чем дать свое согласие, испытуемый должен получить достоверные сведения о сути исследования. Основным источником информации об исследовании для потенциального испытуемого служит «бланк информирован-

¹ Исключением из этого правила являются некоторые формы обсервативного исследования, особенно те, которые проводятся в естественных условиях. Возьмем, к примеру, обсервативное исследование, направленное на изучение процесса принятия решения о покупке в супермаркете, или на изучение того, какие места предпочитают занимать зрители в кинозале. Вряд ли какой-нибудь наблюдательный совет будет настаивать на получении информированного согласия лиц, данные которых будут учитываться в таких исследованиях. Отсутствие согласия можно оправдать несколькими обстоятельствами: поведение испытуемых естественно и на него не влияет факт его изучения; поведение безобидно и не является ни постыдным, ни компрометирующим; испытуемые сохраняют свою анонимность и уверены в этом.

ного согласия», подписание которого свидетельствует о готовности принять участие в исследовании. Цель составления этого документа — донесение в доступной форме, без использования специальных терминов, всех сведений об исследовании, которые необходимы для взвешенного решения о том, принимать ли предложение исследователя. Все наблюдательные советы с особым вниманием относятся к бланкам согласия, которые намеревается использовать исследователь, и все наблюдательные советы сверяют их с действующими нормативными требованиями к этим документам.

Несмотря на наличие четких норм, иногда бланки согласия не в полной мере выполняют свою функцию информирования. Это связано с рядом обстоятельств. Исследователь может столкнуться с трудностями при составлении текста «на языке, доступном для понимания испытуемыми» (American Psychological Association, p. 1608). Большинство психологов имеют довольно скромный опыт составления текстов, адресованных непсихологам, и не все успешно справляются с этой работой. В одном из обзоров наблюдательных советов (Tannenbaum & Cooke, 1977) была произведена оценка доступности изложения в бланках для согласия. Было установлено семь уровней доступности, от «очень легкого» (уровень комиксов) до «очень трудного» (уровень научной и специальной литературы). В обзоре было обнаружено, что 55 % бланков для согласия попадают в категорию «трудных», категорию, соответствующую учебным и академическим работам; 85 % всех бланков попали в категорию трех высших уровней, «умеренно трудных», «трудных» и «очень трудных»; и только 2 % были отнесены к первым трем категориям, «очень легких», «легких» и «умеренно легких». Вывод, следующий из этих данных, очевиден: бланки для согласия не могут выполнять свою функцию информирования, если изложенные в них сведения находятся на уровне, недоступном для понимания многими людьми.

Кроме того, бланки согласия не в полной мере выполняют свое назначение, если часть информации об исследовании намеренно замалчивается. Характер некоторых исследований требует, чтобы испытуемые не знали заранее о цели некоторых процедур. Иногда сокрытие информации принимает форму **неполного раскрытия** — просто утаивания от испытуемых сути некоторых аспектов исследования. Вспомним, к примеру, описанное в главе 5 исследование Хорка и Фэрроу (Horka & Farrow, 1970), в котором изучалась коммуникация между испытуемыми. Валидность такого исследования может сохраняться, только если детям не будет известен ни истинный предмет исследования, которым является коммуникация между испытуемыми, ни факт замены правильного ответа. В других видах исследования отсутствие полного информирования принимает форму активного **обмана**: не просто предоставления неполной информации, а намеренное введение испытуемых в заблуждение относительно некоторых аспектов исследования. Описанное в главе 4 исследование Либберта и Барона (Liebert & Baron, 1972) — наглядный пример. В этом исследовании не было никакого ребенка, игравшего в соседней комнате, и в то же время измерение агрессии основывалось на вере испытуемого в то, что он

либо помогает, либо причиняет боль другому ребенку. В психологии развития ряд исследовательских процедур требует использования либо обмана, либо (чаще) сокрытия части информации. В качестве упражнения было бы полезно, знакомясь с различными исследовательскими подходами, описанными в заключительных главах этой книги, попытаться определить, какие из процедур требуют использования обмана или неполного раскрытия.

Неудивительно, что неполное раскрытие с этической точки зрения представляется более приемлемым, чем активный обман. Тем не менее и неполное раскрытие, и обман подразумевают нарушение принципа информированного согласия, поскольку испытуемые не получают полной информации об исследовании, прежде чем дать свое согласие на участие. В каких случаях использование подобных процедур оправданно?

Из тех норм этики, о которых говорилось выше, можно вывести ряд практических правил. К обману или неполному раскрытию можно прибегать лишь тогда, когда проблематика исследования имеет большое значение, а другого способа собрать необходимые данные не существует. Исследователь может использовать эти приемы только после того, как независимый наблюдательный совет согласится с тем, что их использование оправданно. Хотя испытуемым нельзя рассказать об исследовании все заранее, они должны узнать как можно больше, за исключением, естественно, того, что впоследствии может исказить их ответы. К фактам, о которых испытуемые должны быть осведомлены обязательно, относится риск, связанный с участием в исследовании. По завершении обследования испытуемые должны получить исчерпывающую информацию обо всех целях и экспериментальных процедурах, в отношении которых они находились в неведении. Наконец, и это самое главное, и наблюдательный совет, и исследователь должны быть убеждены, что использование неполного сокрытия или обмана не причинит испытуемым никакого вреда.

Как это всегда бывает, декларировать общие принципы намного легче, чем воплощать их на практике в реальных жизненных условиях. В следующем разделе, обсуждая исследования, которые могут нанести вред испытуемому, я еще вернусь к этим вопросам, в частности к проблеме использования обмана.

Обратимся теперь ко второму компоненту термина «информированное согласие», к понятию «согласие». Чтобы иметь значение, согласие должно подразумевать нечто большее, чем подпись в бланке для согласия: согласие должно быть добровольным. Понятие «добровольность» не означает, что люди должны сами обращаться к исследователям с просьбой разрешить им стать испытуемыми; не означает оно и того, что исследователи должны отказываться от любых попыток убедить потенциальных испытуемых принять участие в исследовании. Однако это означает, что испытуемых нельзя заманивать в исследование ложными обещаниями или принуждать принять участие посредством использования власти. Это означает также, что испытуемых нельзя заставлять продолжать участие в исследовании, если они недвусмысленно выразили желание из него выйти.

Проблема принуждения стоит особенно остро в тех случаях, когда испытуемый и экспериментатор находятся в неравном положении. Яркий пример — когда испытуемые это дети. Еще один пример — отношения между экспериментатором родителями ребенка, особенно когда экспериментатор связан (или родителям кажется, что связан) с учреждением, контролирующим ребенка, например с больницей или школой. В таких случаях родители могут не осознавать, что имеют право отказаться разрешить ребенку участвовать в исследовании. Каковы бы ни были конкретные обстоятельства, исследователь должен найти равновесие между задачей привлечения максимального количества испытуемых и задачей обеспечения истинно добровольного участия испытуемых. Как правило, исследователи уверены в научном значении и этической безупречности того, что они делают; кроме того, они обычно убеждены в том, что их исследование принесет общественную пользу, либо через помощь непосредственно испытуемым, либо перспективу помощи людям, имеющим те же проблемы, что и испытуемые. В этих обстоятельствах исследователю совершенно необходимо распределить свои усилия так, чтобы сформировать и сохранить достаточно большую и в то же время репрезентативную выборку. Однако важно, чтобы его усилия не привели к созданию ошибочного впечатления репрезентативности или к принуждению: ни при привлечении испытуемых, ни в процессе исследования при попытке убедить испытуемых не выходить из исследования.

Теперь рассмотрим, в чем состоит суть информированного согласия при работе с детьми. Дети, ни с юридической, ни с психологической точки зрения, не уполномочены давать информированное согласие на свое участие в исследовании. Следовательно, необходимо обращаться к некоему третьему лицу. В подавляющем большинстве случаев согласие дает родитель (родители) ребенка. Как правило, достаточно согласия одного из родителей ребенка, однако бывает и так, что наблюдательный совет требует получения согласия обоих родителей. В редких случаях наблюдательный совет решает, что в родительском разрешении нет необходимости, точно так же, как это бывает в ситуациях, когда выносится решение об отсутствии необходимости получения информированного согласия взрослых испытуемых. Однако даже тогда, когда от исследователя могут потребовать разрешение школы, где обучается ребенок, такое разрешение может быть условием, при котором снимается требование получения родительского согласия. Бывают также ситуации, когда отсутствует возможность связаться с родителями ребенка — к примеру, когда ребенок живет в приюте или на попечении. Однако принцип остается тем же: должно быть получено разрешение того лица, которое несет юридическую ответственность за ребенка и способно понять, в чем будет заключаться участие ребенка в исследовании.

Так же как и для бланков информированного согласия, институциональные наблюдательные советы выдвигают требования к письменным обращениям к родителям. Во вставке 8.2 представлены (в несколько сокращенном виде) нормы по составлению обращения к родителям, разработанные Институциональным наблюдательным советом штата Флорида.

Вставка 8.2

Требования к информированному разрешению родителей (Институциональный наблюдательный совет университета штата Флорида)

А) Представьтесь и уточните, каким образом вы связаны с университетом штата Флорида.

Б) Укажите на необходимость проведения исследования; объясните, что потребуется от испытуемых. Уточните, кто будет проводить процедуры (например, вы, учитель, ваш супервизор, коллега). Убедитесь, что язык, которым вы пользуетесь, доступен для понимания тем, кому предназначен этот документ. Избегайте специальных терминов и профессионального жаргона.

В) Укажите приблизительное время, которое займут исследовательские процедуры (количество дней, минут и т. д.)

Г) Если процедуры будут производиться в учебное время, уточните, чем именно в это время будут заниматься дети. Проинформируйте родителя (родителей), будет ли ребенок снят с урока, и если это так, будет ли у ребенка возможность восполнить этот пропуск.

Д) Если есть необходимость, отметьте, что участие или неучастие в исследовании не отразится на школьных отметках ребенка и его учебном статусе.

Е) Если вы собираетесь обследовать весь класс или всю группу, уточните, чем во время исследования будут заниматься те, кто не получил разрешения родителей.

Ж) Включите описание любых нежелательных последствий и неудобств. Укажите, какую пользу ребенок может извлечь из участия в исследовании, в том числе приобретение новых знаний. Если участие в исследовании не подразумевает никаких выгод для ребенка, отметьте и это.

З) Укажите, каким образом будет сохранена анонимность ребенка (анонимность бланков с ответами, использование кода и т. д.) и, если необходимо, заверьте в соблюдении конфиденциальности. (Например: «Личные результаты учащегося будут сохранены в тайне посредством использования системы кодирования; гласности будут преданы только групповые результаты».) Уточните, получит ли кто-либо еще доступ к данным, и если это так, объясните, к какого рода данным, групповым или индивидуальным, и кто (например, родители, учителя, школьный совет).

И) Укажите, будет ли использовано какое-либо поощрение за участие в исследовании. Если предлагается денежное вознаграждение, укажите сумму. Если предполагается другая форма вознаграждения, укажите ее (например, наклейка, карандаш).

К) Проинформируйте родителей об их (и ребенка) праве в любое время отказать в согласии на участие в исследовании или на использование данных о ребенке без каких-либо последствий или какого-либо ущерба. **Не** ставьте условий на право отказа.

Л) Если будут использованы опросники, анкеты или интервью, отметьте, что ребенок не обязан отвечать на те вопросы, на которые не пожелает отвечать. **Не** делайте никаких оговорок относительно этого права ребенка.

М) Если частью исследовательского процесса является запись, укажите это при получении информированного согласия родителей. Уточните, будет использоваться аудио- или видеозапись, кто будет иметь доступ к пленкам и как в дальнейшем распорядятся этими материалами.

Н) Предложите ответить на возникшие вопросы и предусмотрите в бланке информированного согласия средства обратной связи.

О) Укажите, что с вопросами о правах испытуемых можно обращаться в управление Наблюдательного совета университета Флориды.

П) Бланк информированного согласия родителей должен оканчиваться следующим или равноценным утверждением: «Я прочел (прочла) представленное выше описание исследовательской процедуры. Я согласен (согласна) разрешить своему ребенку, _____, участвовать в исследовании (название или краткое описание) доктора _____, и я получил(а) копию этого описания».

Требование получения родительского разрешения практически на все исследования с детьми было введено недавно. Еще одно достижение — значительно большее внимание к праву ребенка решать, принимать или не принимать участие в исследовании. В отличие от процедуры получения согласия взрослых, получение согласия детей, как правило, производится в устной форме: исследователь кратко описывает исследование, а затем спрашивает ребенка, хотел бы он принять в нем участие или нет. Положительный ответ ребенка называют «разрешением» (*assent*), в отличие от юридически значимого «согласия» (*consent*), которое дают взрослые.

Убедить ребенка принять участие в исследовании не всегда легко. Подход в этом деле зависит от возраста ребенка и от характера исследования. Хотя точных данных на этот счет нет, по-видимому, не будет преувеличением заметить, что обычно исследователи рассказывают детям об исследовании значительно меньше, чем взрослым. Это оправдано в том смысле, что ребенку труднее понять цели исследования. Еще одно основание заключается в том, что многие взрослые уже дали свое согласие на участие ребенка в исследовательском проекте. Тем не менее, чтобы иметь хоть какое-то значение, разрешение ребенка должно подразумевать получение ребенком некоторых оснований для принятия решения, идти или не идти с экспериментатором. За исключением случаев работы с очень маленькими детьми, имеет смысл донести до ребенка следующее: общую картину того, что будет происходить («поиграем в игры на запоминание»), где будет проводиться исследование («в кабинете у мистера Смита»), сколько людей будет задействовано («только ты и я»), сколько потребуется времени («это займет двадцать минут»), делали или будут делать то же самое другие дети («многие ребята из твоего класса тоже будут делать это») и, наконец, возможность сказать да или нет («Хочешь прийти?»). Чем старше испытуемый, тем ближе должно быть описание процедуры к тому, которое получает взрослый, давая информированное согласие (см. также Tymchuk, 1991).

Абрамович, Фридман, Тоден и Николитч в своем отчете (Abramovich, Freedman, Thoden, & Nicolich, 1991) приводят интересные данные о том, как дети понимают процедуру получения разрешения на участие в исследовании. Сначала авторы провели стандартную процедуру получения разрешения, в ходе которой были разъяснены права ребенка как испытуемого, а затем проверяли, насколько хорошо их 5–12-летние испытуемые поняли свои права (например: «Что произойдет, если ты больше не захочешь принимать участие в исследовании?»; «Кто узнает о том, что ты делал во время исследования?»). Результаты оказались неоднозначными. С одной стороны, большинство детей понимало, о чем их попросят экспериментаторы и что они смогут по собственному желанию отказаться от дальнейшего участия в исследовании. Однако, с другой стороны, они не были уверены в том, что могут сразу отказаться от участия, особенно если родители уже дали свое согласие. Кроме того, несмотря на все заверения исследователя в том, что результаты станут известны только коллегам по исследованию, большинство детей полагало, что об их результатах узнают родители. Как отмечают Абрамович и соавторы, для детей более привычна ситуация, когда их родители знают о том, что происходит с ними в школе; поэтому им трудно понять, что собранные в школе исследовательские данные носят конфиденциальный характер.

Ненанесение ущерба

Несомненно, важнейшим из прав, которыми обладают испытуемые, является право на ненанесение ущерба в ходе исследования. Соответственно основная обязанность — следить, чтобы ни одному испытуемому не был причинен вред.

Какой вред может быть причинен испытуемому? В медицинских исследованиях иногда возникает риск физического вреда. Новый вид лечения, к примеру, может как принести пользу, так и привести к нежелательным последствиям, и до окончания исследования такой риск исключать нельзя. Именно в таких исследованиях информированное согласие, равно как и предварительная оценка наблюдательного совета, приобретают особое значение. В действительности, появление первых наблюдательных советов было связано как раз с необходимостью контролировать исследования в области медицины. В психологии эти советы появились несколько позже.

В психологическом исследовании опасности причинить физический вред практически нет. Здесь наибольшее значение имеет возможность психологического вреда. Психологический вред — менее определенное понятие, чем физический вред и, будучи таковым, является предметом многочисленных споров как при обсуждении этики вообще, так и при оценке конкретных исследовательских проектов. Начнем с общих замечаний, а затем перейдем к рассмотрению исследований, в ходе которых может быть причинен психологический вред.

Понятие психологического вреда не означает, что в ходе исследования испытуемый не должен испытывать никакого дискомфорта. Испытуемому, к примеру, может наскучить выполнение монотонного задания, он может пожалеть, что занимается этим делом, может счесть некоторые элементы процедуры фрустрирующими или неприятными, может не справиться с заданием, на успех в выполнении которого он рассчитывал, и т. д. Конечно, такие переживания следует свести к минимуму, особенно если исследование не направлено на изучение скуки или фрустрации. Однако сам по себе тот факт, что испытуемый предпочел бы оказаться в другой обстановке, не означает, что исследование наносит психологический вред. Нужно ответить на два вопроса. Насколько резко опыт, получаемый в ходе исследования, отличается от того, что для испытуемого привычно? Есть ли вероятность того, что последствия будут стойкими, а не преходящими?

Рассмотрим исследование, в котором изучается способность школьников к решению задач. Детям предлагается ряд задач; одни дети справляются с ними лучше других; и у всех детей не получается хотя бы одна задача. В таком виде исследование представляется весьма похожим на то, что обычно происходит в школе, и нет никаких оснований полагать, что негативные последствия неудачи при решении задач выйдут за пределы ситуации исследования. Предположим, однако, что одна из целей исследования — определить влияние тревоги на процесс решения задач. Для повышения тревоги половине детей сообщают, что задачи специально подобраны для проверки возможностей успешного обучения в следующем году. В этом случае неудачи при решении задач вполне могут привести к серьезным негативным реакциям, которые сохранятся и после исследования.

Этот пример иллюстрирует один из типов экспериментальных манипуляций, которые могут нанести психологический ущерб, манипуляций, вызывающих у

испытываемого негативные переживания или создающие негативный Я-образ. К этой категории относятся искусственно вызванные переживания неудачи, особенно при выполнении заданий, которые представляются как задания, имеющие большое значение. Сюда также входят приемы для изучения наказания, о которых мы упомянем в главе 12, а также манипуляции, заставляющие испытуемого вести себя неэтично или нарушать нормы. Яркий пример — мошенничество. В исследованиях, направленных на изучение этого явления, детей (или по крайней мере некоторых детей) побуждают сделать что-то, что, как им известно, противоречит правилам и при раскрытии вызовет порицание. Даже если ребенка заставляют поверить, что о его мошенничестве никто не узнал, у него может возникнуть чувство тревоги или вины. Если ребенка как-то не переубедить, эти чувства могут закрепиться.

Сказанное выше подсказывает способ разрешения подобных проблем. Почему бы не использовать период после обследования для снятия неприятных ощущений, возникших в ходе исследования? В исследовании с решением задач, к примеру, экспериментатор мог бы объяснить детям, что результаты тестирования в действительности ничего не говорят об успеваемости и что все дети какие-то из заданий выполнили неправильно. В исследовании с мошенничеством экспериментатор мог бы подчеркнуть, что оно не причинило никому вреда и что многие дети вели себя точно так же. Такое раскрытие истинных целей исследования после свершившегося факта называется **дебрифингом**. В ходе дебрифинга решаются две задачи: прояснение действительных целей экспериментальных манипуляций и создание всех условий для того, чтобы испытуемый не чувствовал себя неловко из-за своего поведения.

Дебрифинг рекомендуется проводить во всех случаях, когда изначально от испытуемых скрывается часть информации. Если во время исследования вызываются негативные чувства, тогда проведение дебрифинга приобретает еще большее значение. Однако с детьми проведение дебрифинга бывает затруднено и подчас создает еще большие проблемы, а не решает их. Ребенок не обязательно будет чувствовать себя лучше, если экспериментатор завершит обследование объяснением, что производилось изучение мошенничества и что все это время за ним наблюдали. В конечном счете ребенок действительно провинился, и теперь он даже не может утешиться тем, что сделал это тайком. Неизвестно также, в какой мере маленькие дети могут понять объяснения истинных целей исследования и насколько это может им помочь. Экспериментатор, пытающийся обосновать приемы, использовавшиеся в исследовании, может лишь вызывать у 5-летнего ребенка смущение или тревогу. Что если взрослые могут смотреть через любые зеркала?! В таких обстоятельствах лучше вообще отказаться от дебрифинга.

Дебрифинг может иметь и еще одно неприятное последствие. По определению, дебрифинг проводится только после того, как экспериментатор тем или иным образом ввел испытуемого в заблуждение. В ряде случаев, «ввел в заблуждение» — это всего лишь эвфемизм; по сути, экспериментатор обманывает испытуемого. Таким образом, дебрифинг развенчивает ложь. Однако может случиться, что главный вывод, который сделает для себя испытуемый, будет состоять в том, что исследователям (или психологам, или взрослым вообще) нельзя доверять. Дети особенно подвержены негативным реакциям на ложь взрослого.

Сказанное выше напоминает убийство гонца, принесшего дурную весть. Ясно, что главный обвиняемый здесь не дебрифинг, а изначальный обман. Очевидный способ избежать проблем, сопутствующих обману испытуемых, — вообще не лгать. Однако некоторые вопросы (например, мошенничество) без обмана изучать весьма затруднительно. Если все требования, описанные в предыдущем разделе, соблюдены, большинство исследователей согласится, что использование обмана иногда оправдано (однако не все исследователи; ср. Baumrind, 1985). Именно в этих случаях исследователь — и институциональный наблюдательный совет — должны решить, смягчит ли дебрифинг проблемы, вызванные обманом, или лишь обострит их.

Пример исследования мошенничества иллюстрирует еще один тип психологического вреда. Мы рассмотрели два типа возможных негативных последствий участия в исследовании с мошенничеством: чувство вины из-за совершения проступка и потеря доверия при раскрытии обмана. Однако есть и третье следствие, в некотором смысле, более прямое. Допустим, ребенок не испытывает угрызений совести из-за своего поведения, а, наоборот, доволен тем, что открыл новый способ, как перехитрить взрослого. Предполагается, что конечная цель исследования мошенничества состоит в выявлении того, что может его предотвратить. Однако непосредственным результатом нашего исследования может оказаться поощрение поведения, которые мы хотим пресечь.

Это звучит даже еще более убедительно в контексте другого примера. Рассмотрим вопрос, который уже не раз обсуждался в предыдущих главах: связь между просмотром телепрограмм с эпизодами насилия и агрессивным поведением. Все подходы к экспериментальному изучению этой проблематики предполагают намеренный показ ребенку телематериалов со сценами насилия, который, по мнению исследователей, должен повысить уровень агрессии. Часто просмотр таких материалов действительно повышает уровень агрессии в общении с другими детьми. Таким образом, исследование стимулирует социально нежелательное поведение, поведение, которое, кроме того, направлено на невинных жертв. Можно ли оправдать такое исследование?

Если такие исследования вообще можно оправдать, то оправдывающие их обстоятельства близки к тем, о которых говорилось выше. Агрессия — в высшей степени важная тема, как с научной, так и социальной точки зрения. Понимание ее причин, включая и просмотр агрессивных телепрограмм, имеет особое значение. Для точного определения причинно-следственных связей требуется экспериментальный подход, что не раз подчеркивалось на страницах этой книги. Опыт, который моделируется в таких условиях — несколько часов просмотра телепрограмм с эпизодами насилия, стычки на игровой площадке, — не слишком отличается от повседневного опыта детей. Если учесть непродолжительность экспериментального опыта и множественность детерминант агрессии, любой стойкий эффект экспериментальных манипуляций кажется маловероятным. Иными словами, при таком исследовании риск ничтожен, а потенциальная польза велика.

Хотя такое оправдание вполне приемлемо, исследователь должен прибегать к нему с большой осторожностью. Экспериментальные процедуры, провоцирующие нежелательное поведение, должны использоваться только в случаях крайней не-

обходимости. Общий принцип звучит следующим образом: не следует поощрять такое поведение детей, которое нам (и возможно им) кажется нежелательным.

Последняя, проблематичная с точки зрения этики ситуация в определенном смысле противоположна ситуации, описанной выше. Только что мы рассмотрели случаи, когда исследователь намеренно создает условия для нежелательного поведения испытуемых, например, провоцирует агрессию. Однако бывают и такие ситуации, когда исследователь намеренно не создает условий для желательных результатов. Эту проблему, как правило, рассматривают под рубрикой **отказ в потенциально благотворном воздействии**, то есть невозможность сделать доступным некое, в перспективе полезное для испытуемого, экспериментальное воздействие.

Впервые проблема отказа в экспериментальном воздействии встала в медицинских исследованиях. Допустим, мы разработали новый вид лекарственной терапии, который, по нашему мнению, должен оказаться эффективным в борьбе с определенной формой рака. Чтобы проверить свою гипотезу, мы должны применить нашу терапию на выборке людей, страдающих этим заболеванием. Нам также нужна контрольная группа, которая не будет проходить этот вид терапии и сравнение с которой даст возможность оценить эффективность разработанной программы лечения. В одних случаях для контроля оказывается достаточно использовать уже имеющиеся данные о естественном ходе заболевания. Однако в некоторых случаях точное сравнение можно произвести только при проверке двух изначально эквивалентных групп: экспериментальной группы, которая пройдет разработанный курс лечения, и контрольной группы, которой мы откажем в этой возможности. Таким образом, у нас будет точный критерий эффективности нашей терапии. Однако как оправдать отказ в потенциально благотворном воздействии людям, которые в нем нуждаются?

Рассмотрим аналогичный психологический пример. Последние 30 лет разрабатывалось множество программ вмешательства в дошкольном возрасте, направленных на улучшение подготовленности к школе маленьких детей. Как правило, такие программы ориентированы на детей, которых по разным причинам (например, низкий социально-экономический статус) считают входящими в «группу риска» школьной неуспеваемости. Для определения эффективности любой такой программы требуются две группы детей: экспериментальная группа, которая будет проходить программу, и равноценная ей контрольная группа детей, которые не будут проходить эту программу. Однако мы знаем, что без некоторого вмешательства многие дети из контрольной группы не смогут успешно обучаться. Как оправдать отказ во включении их в нашу программу?

Можно предложить ряд оправданий намеренного отказа в потенциально благотворном воздействии. Во-первых, мы не знаем, эффективно ли наше воздействие; если бы мы это знали наверняка, не было бы необходимости проводить экспериментальную проверку. Только имея возможность экспериментальной проверки собственных идей, мы сможем доказать ценность нашей программы, и это — доказательство, которое в дальнейшем может принести пользу значительно большему количеству людей. Во-вторых, у исследователя редко бывает достаточно ресурсов, чтобы предложить воздействие всем людям или большей части людей, которым

оно могло бы принести пользу. Учитывая это ограничение, разумнее всего использовать подход, позволяющий одновременно принести пользу максимально возможному количеству людей, при этом осуществляя научную проверку эффективности экспериментального воздействия. Основной этический принцип разделения на терапевтическую и контрольную группы состоит в использовании случайного причисления, без учета таких факторов, как социально-экономический статус, раса или дружеские отношения испытуемого с исследователем. Всегда, когда есть возможность, контрольной группе следует предложить оказание терапевтической или хотя бы какой-то иной формы помощи по завершении экспериментальной фазы исследования. Наконец, иногда есть возможность сопоставления не экспериментальной и контрольной групп, а двух или более видов экспериментального воздействия. В этом случае все испытуемые имеют возможность извлечь пользу из исследования.

Конфиденциальность

Последнее из основных прав испытуемого — это право на **конфиденциальность**. Использование информации, полученной в ходе исследования, должно быть ограничено научными целями, о чем испытуемому нужно сообщить до получения информированного согласия. Такая информация не должна быть доступной в том случае, если это может смутить испытуемого или причинить ему вред.

В процессе исследования проблема конфиденциальности становится актуальной два раза. Во-первых, во время сбора данных и помещения их на хранение. Понятно, что если на бланках с данными будут стоять фамилии испытуемых, всегда существует вероятность, что посторонние лица увидят эти бланки и узнают то, что не подлежит разглашению. Очевидным выходом из этой ситуации является документирование и хранение данных с использованием кодов вместо имен. Если когда-то может понадобиться установить связь между данными и именами, отдельно от данных следует хранить список перевода кодов. Разглашение информации может произойти и тогда, когда в своих разговорах исследователь упоминает фамилии испытуемых или как-то иначе дает понять, о ком идет речь. Естественное решение этой проблемы — воздерживаться от таких разговоров.

Второй раз вопрос конфиденциальности становится актуальным при обнародовании результатов. Цель обнародования, конечно же, состоит в сообщении того, что было сделано и достигнуто, максимально широкой аудитории. Однако, сообщая об этом, важно обеспечить анонимность испытуемых. В большинстве психологических отчетов сделать это нетрудно. Обследуются большие выборки, а результаты анализируются с точки зрения группового среднего и межгрупповых различий. Однако есть возможность установить личность каждого из испытуемых — к примеру, в отчетах об исследовании отдельных случаев, небольших или весьма своеобразных выборок. В этих условиях для обеспечения конфиденциальности может потребоваться использование специальных приемов (например, введение псевдонимов, опускание географической информации).

Как следует из сказанного выше, в большинстве исследовательских проектов никаких особых препятствий к сохранению конфиденциальности нет. Все, что необходимо делать исследователю, это избегать небрежности в разговорах и обеспе-

чить недоступность бланков с данными для посторонних лиц. Рассмотрим теперь две ситуации, в которых сохранить конфиденциальность несколько сложнее.

Исследователей, изучающих развитие ребенка, часто просят сообщить об успешности выполнения ребенком использовавшихся тестов. Учитель может захотеть узнать, как тот или иной ребенок из его класса справился с когнитивным тестом. Родители могут захотеть узнать, что показали использовавшиеся исследователем методики о развитии их собственного ребенка. Это довольно щекотливая ситуация. С одной стороны, исследователь получил возможность обследовать детей только благодаря любезности родителя или учителя. Отказ поделиться информацией может показаться неразумным шагом и может препятствовать дальнейшему сотрудничеству. С другой стороны, считается, что ребенок, который выступает в качестве объекта исследования, обладает тем же правом на конфиденциальность, что и любой другой испытуемый. В некоторых исследованиях (например, восприятия детьми методов воспитания, используемых их родителями) важна уверенность ребенка в том, что о его ответах никто не узнает.

Реагируя на просьбы об обратной связи со стороны родителей или учителей, следует руководствоваться некоторыми общими правилами. Первое правило — проясните с самого начала, какая информация будет или не будет предоставляться. Если результаты ребенка не будут подлежать разглашению, это должно быть отражено в бланке для согласия, который подписывают родители. Второе правило — студенты, занимающиеся тестированием, должны быть особенно осторожны в вопросе предоставления обратной связи учителю или родителю, настаивающему на получении информации частного характера; нужно направить их к руководителю. Третье правило — старайтесь давать позитивную оценку, в то же время не прибегая ко лжи. Всегда можно сказать что-нибудь положительное о результатах ребенка, хотя, конечно, не стоит рисовать неоправданно радужную картину. Негативную оценку следует давать крайне осторожно. Последнее правило: указывайте на то, что диагностические возможности ваших методик не беспредельны. В большинстве исследований в области психологии развития цель диагностической оценки индивидуальных различий детей не стоит; а большинство методик, используемых в психологии развития, не предназначены для получения информации об индивидуальных особенностях ребенка, обладающей диагностической и прогностической ценностью. Если учителям и родителям станет известно об ограниченных возможностях тестов, проблем с запросами информации частного характера может и не возникнуть. Заметьте, что сами испытуемые имеют право узнать о том, какие выводы можно сделать на основе их результатов, а какие — нельзя.

Хотя большинство методик не позволяет глубоко изучить конкретного ребенка, из этого правила существуют исключения, и эти исключения представляют собой вторую опасность, угрожающую конфиденциальности. Допустим, в исследовании обнаружилось, что в развитии ребенка есть серьезное нарушение, нарушение, которое представляется столь серьезным, что к нему необходимо привлечь внимание лица, ответственного за благополучие ребенка. При использовании батареи когнитивных тестов может быть получен вызывающий тревогу низкий показатель, столь низкий, что исследователь сочтет необходимым оповестить об этом

школьное руководство. При использовании набора личностных методик может выявиться странность в ответах, свидетельствующая о глубоком личностном расстройстве. В исследовании, даже случайно, может обнаружиться нечто, указывающее на жестокое обращение с ребенком. Как и во всех исследованиях, информация здесь носит конфиденциальный характер. Что же обязан сделать исследователь?

Ответ на этот вопрос дает кодекс этики, разработанный Обществом исследований в области детского развития и представленный во вставке 8.1. Девятый принцип кодекса гласит: «Если в ходе исследования исследователь узнает о чем-либо, что угрожает благополучию ребенка, он должен обсудить этот вопрос с родителями или опекунами и специалистами в данной области, с тем чтобы оказать ребенку необходимую помощь». Из этого положения, каким бы абстрактным оно ни было, можно сделать два вывода. Во-первых, принцип конфиденциальности не является чем-то незбылемым. Конечной целью всегда является благополучие ребенка, а в некоторых случаях для обеспечения благополучия испытуемого необходимо нарушить конфиденциальность. Нет необходимости подчеркивать, какую серьезную ответственность берет на себя исследователь, сообщая родителям, что их ребенок, вероятно, имеет психические нарушения, или информируя правоохранительные органы о том, что родитель, вероятно, жестоко обращается с ребенком. Цена ошибочных действий или бездействия крайне высока. Прежде чем решиться на какие-то шаги, исследователь должен получить квалифицированный совет.

Резюме

Любое исследование должно отвечать двум требованиям: научности и этической безупречности. В этой главе рассматриваются этические принципы, которыми необходимо руководствоваться при проведении исследования в области психологии развития.

Глава начинается с изложения общих норм, в соответствии с которыми должен действовать исследователь в области психологии развития. За последние годы было разработано множество разнообразных кодексов этического поведения, к которым может обратиться исследователь, оценивая этичность собственного исследования. Наиболее ценный источник для специалиста в области психологии развития — нормативы Общества исследований в области детского развития, представленные в данной главе.

На местном уровне следование этическим принципам контролирует организация, которая называется институциональный наблюдательный совет. Все исследования с людьми должны предварительно получить одобрение институционального наблюдательного совета. Требование этого одобрения отражает базовый принцип исследования с людьми: необходимость независимой оценки соответствия исследовательского проекта этическим нормам. В исследованиях с детьми дополнительным охранительным барьером является требование разрешения на участие ребенка в исследовании со стороны родителей, а в некоторых случаях — школьной администрации. Однако, как подчеркивается, одобрение других людей необходимое, но не достаточное условие при оценке этичности. Ответственность за следование этическим стандартам в конечном счете лежит на самом исследователе.

За рассмотрением норм и правил следует обсуждение основных прав испытуемого. Одно из них — право дать информированное согласие прежде чем принять участие в исследовании. Испытуемым нужно заранее объяснить, в чем суть исследовательского проекта, чтобы они смогли свободно принять решение, принимать или не принимать участие в исследовании. Анализируются разнообразные обстоятельства, препятствующие предоставлению испытуемому полной информации об исследовании. В одних случаях исследователю не удастся разъяснить цели и порядок проведения своего исследования на языке, соответствующем возможностям испытуемых. В других случаях исследователь намеренно утаивает информацию об определенных аспектах своего исследования либо в форме неполного информирования, либо в форме обмана. При изучении некоторых вопросов такая закрытость просто необходима, тем не менее важно, чтобы испытуемые узнали как можно больше, включая правду о любом риске. При изучении детей могут возникнуть особые трудности с получением согласия. Родители или другие ответственные за ребенка лица должны быть проинформированы об исследовании и дать свое разрешение на участие в нем ребенка. Кроме того, ребенку необходимо доступным для него языком описать исследование с тем, чтобы он имел возможность дать или не дать свое разрешение на обследование.

Второе из основных прав испытуемого — право на ненанесение ущерба. В психологическом исследовании особое значение имеет психологический вред, спорное понятие, которому довольно трудно дать четкое определение. При оценке вероятности причинения вреда экспериментальными манипуляциями важны два момента. Получает ли испытуемый в ходе исследования опыт, резко отличающийся от обычного для него опыта? Есть ли вероятность того, что последствия экспериментального воздействия окажутся стойкими? К манипуляциям, которые могут нанести вред, относятся те, которые создают негативный Я-образ, вызывают неприятные чувства (например, враждебность, тревогу) и провоцируют нежелательное поведение. Всегда, когда используются такие манипуляции, как правило, есть смысл по окончании обследования провести дебрифинг. Еще один тип ситуаций, проблематичных с этической точки зрения, — отказ в потенциально благотворном воздействии: людям, которые могли бы извлечь пользу из экспериментального воздействия, в нем отказывается. Если есть возможность, нужно сделать так, чтобы из исследования смогли извлечь пользу и испытуемые из контрольной группы; например, предложить им альтернативную форму терапевтического воздействия или ту же экспериментальную форму воздействия, но по окончании исследования.

Последнее из основных прав, анализируемых в этой главе, — право на конфиденциальность. Данные, полученные в ходе исследования, носят конфиденциальный характер и должны быть недоступны в той мере, в которой они могут нанести испытуемому вред. Поэтому важно сохранять анонимность испытуемых, как при сборе данных, так и при обнародовании результатов. Сохранить конфиденциальность представляет особые трудности в случаях, когда информацию об индивидуальных результатах запрашивают учителя или родители; в этой главе предложены некоторые рекомендации о том, как следует реагировать на подобные просьбы. Еще более затруднительная ситуация — когда исследователь обнаруживает признаки серьезного неблагополучия ребенка (например, признаки жестокого обращения с

ребенком). В таких случаях конфиденциальность можно нарушить, поскольку на первом месте всегда стоит благополучие ребенка.

Упражнения

Спланируйте (по крайней мере в общих чертах) исследование по той проблематике детской психологии, которая вас более всего интересует. Представьте, что вы подаете свой проект на рассмотрение институциональному наблюдательному совету, и составьте обращение к родителям, которое вы бы включили в свою заявку. Если есть возможность, попросите преподавателей оценить ваше обращение.

В статье, посвященной уязвимости в исследовании («Vulnerability in Research: A Developmental Perspective on Research Risk»), Росс Томпсон (Ross Thompson) делает три зарисовки исследований в области психологии развития, в которых встают вопросы этической безупречности. Прочитайте зарисовки Томпсона (источник — *Child Development*, 1990, 61, 1–16) и дайте собственную оценку этичности этих исследований. Дайте свой ответ с позиции двух людей — члена институционального наблюдательного совета и родителя, которого просят позволить своему ребенку принять участие в исследовании.

Глава 9

Публикация

В главе 1 в общих чертах были описаны восемь этапов исследования. Сейчас мы подошли к последнему из этих этапов — сообщению о своей работе другим людям.

Ученые сообщают о результатах своих исследований двумя способами. Первый из них — устный отчет. Этот метод сообщения используется на профессиональных конференциях; это также, разумеется, происходит в процессе преподавания. Ни одному студенту, прослушивающему дюжину лекций в неделю, не нужно объяснять, что устные сообщения различаются по степени ясности и удачности.

Второй метод — письменное сообщение, обычно в форме публикации в одном из множества специализированных журналов. Именно о данном методе пойдет речь в этой главе. Основной вопрос поэтому — как написать отчет для публикации в журнале АПА (Американской психологической ассоциации — *American Psychological Association*). Однако многие из сделанных рекомендаций имеют более широкое применение. Так, большая часть сказанного имеет отношение и к подготовке устного доклада, а также написанию любых работ, а не только отчетов о психологических исследованиях.

Как всегда, можно обратиться к ряду других источников. Незаменимый источник информации о написании психологических работ — «*Publication Manual of the American Psychological Association*» (APA, 1994). В *Manual* изложены как общие рекомендации, так и конкретные правила написания отдельных элементов рукописи (например, заголовки, подстрочные сноски, список литературы). В этой главе мы будем часто ссылаться на *Manual*. Рекомендации по написанию исследовательских отчетов и другого рода психологических работ (например, рецензий, исследовательских заявок) можно найти в вышедших недавно книгах, адресованных психологами (Parrot, 1994; Rosnow & Rosnow, 1992; Sternberg, 1993). Среди множества хороших книг, посвященных общим вопросам стилистики, классическая работа Странк и Уайт «Основы стиля» (*The Elements of Style*, Strunk & White, 1979). Наконец, ни один читатель, обладающий чувством юмора, не пропустит «Основные принципы подготовки статей для психологических журналов» Харлоу (*Fundamental Principles for Preparing Psychology Journal Articles*, Harlow, 1962).

Структура этой главы такова. Глава начинается с нескольких общих рекомендаций о том, как следует писать психологическую и любую другую работу. Затем

по порядку рассматриваются разделы психологического отчета: «Введение», «Метод», «Результаты» и «Обсуждение». Главу завершает анализ двух проблем, имеющих отношение ко всем авторам: нарушение стандартов АПА и нарушение правил английской грамматики.

Общие замечания

При написании психологической работы необходимо разрешить два основных вопроса. Первый — что сказать, а второй — как это сказать.

В этой книге уже немало говорилось о том, какого рода информацию следует включать в психологическую статью. Я, к примеру, объяснял, что нужно сообщать читателям об испытуемых, о тестирующих, о методах сбора данных. Если говорить в общем, для того чтобы решить, о чем нужно упомянуть, делая отчет об исследовании, необходимо иметь ясное представление о важнейших компонентах исследования, которым были посвящены предыдущие восемь глав. При обсуждении различных разделов исследовательского отчета мы более подробно остановимся на проблеме, связанной с вопросом «что сказать».

Проблемы, связанные с вопросом «как сказать», делятся на две категории. Ряд стилистических норм по существу — условности, то есть произвольные правила представления отдельных аспектов письменной работы. К примеру, нет какого-то «правила» использования подстрочных сносок или составления списка литературы; несомненно, каждый читатель в разных книгах и статьях сталкивался с разными подходами. Однако именно потому, что существует такое разнообразие, представителям научного сообщества необходимо выработать некий единый подход. Поэтому в психологии *есть* правила использования подстрочных сносок, составления списка литературы и ряд других стилистических условностей, которые зафиксированы в *Manual APA*. Говоря словами *Manual* (1994), эти правила «избавляют читателей от отвлекающего внимания разнообразия форм в одной и той же работе и дают возможность сосредоточиться на содержании» (р. XXV).

Я не буду пытаться осветить все стилистические правила, изложенные в «Руководстве», правила, занимающие 348 страниц текста. Однако я попытаюсь сделать две вещи. Во-первых, разъяснить несколько правил, представляющих для студентов особые трудности. Во-вторых — подчеркнуть, как важно приобрести *Manual* и следовать его указаниям при написании психологической работы. С практической точки зрения, лучше всего считать, что в *Manual* отражен *любой* из стилистических вопросов; и действительно, скорее всего, так и есть, начиная с вопроса о ширине полей.

Следование стилистическим нормам АПА — сравнительно простая часть вопроса «как сказать». Более трудная часть имеет отношение не только к психологии. Это общая проблема, как изложить свои мысли ясно и доступно. Нет необходимости говорить, что на этот вопрос нельзя ответить в такой короткой главе как эта (увы, не все так просто!). Однако можно попытаться решить более скромные задачи. В этом разделе сделано несколько общих замечаний о написании научной работы. Затем будут рассмотрены частные стилистические вопросы, представляющие особую трудность для студентов. В этом последнем разделе отражены некоторые специфические нормы, разработанные АПА.

Начну с того, что подчеркну необходимость писать ясно и грамматически правильно. То, что об этом кажущемся очевидным требованием приходится говорить, обусловлено двумя естественными, но тем не менее ошибочными представлениями. Первое — представление о том, что с завершением исследования и получением неких результатов самая важная и интересная часть работы исследователя закончена. Сообщение же о своих результатах — лишь скучное дополнение к настоящей исследовательской работе. Второе — представление о том, что имеет значение только содержание сообщения; «стиль» — это тонкость, которую можно оставить для уроков английского.

Что не учитывается в этих представлениях, так это базовые принципы, о которых говорилось в главе 1: наука — сфера информационного обмена. Чтобы результаты исследования имели хоть какой-то смысл, их нужно сделать доступными для всего научного сообщества. Кроме того, их необходимо представить так, чтобы привлечь к ним внимание и создать условия для их ассимиляции. Письменное изложение поэтому должно заинтересовывать, иначе и без того занятый ученый быстро переключится на что-то другое. Оно должно быть понятным, с описанием всех важнейших подробностей, иначе у читателя не будет возможности оценить значение результатов. Кроме того, оно должно быть аргументированным, с убедительным доказательством значимости проблемы, безупречности методов и валидности выводов. Иными словами, оно должно быть хорошо написано. Чем лучше автор, тем выше вероятность того, что его труд окажет влияние на ту научную область, в которой он работает.

Необходимость писать хорошо подтверждается еще одним, даже еще более очевидным прагматическим обстоятельством. Прежде чем со статьей сможет ознакомиться научная общественность, она должна быть принята для публикации в специализированном журнале. Большинство журналов придирчиво отбирают статьи для публикации, а в лучших журналах уровень отказов приближается к 90 %. У плохо написанной статьи просто мало шансов выйти в свет. Загруженным работой редакторам и рецензентам вряд ли захочется вникать в суть плохо написанной работы, и вполне вероятно, что даже при желании им будет трудно ее понять. Кроме того, поскольку функция исследовательского отчета это сообщение, оценка качества написанного здесь вполне оправдана. Заблуждается тот, кто считает, что при оценке написанного можно или следует разделить стиль и суть.

Осознать необходимость писать хорошо нетрудно; гораздо труднее хорошо написать. Как отмечалось, в этой главе я ограничусь изложением некоторых общих рекомендаций.

Первая рекомендация — читайте психологические работы, прежде чем писать психологические работы. Основная трудность, с которой сталкиваются многие студенты при написании исследовательского отчета, заключается в том, что сами они почти не читали таких отчетов. Как будет подчеркнуто ниже, написание психологической работы ничем кардинально не отличается от написания любой другой работы. Тем не менее есть своего рода чутье — на то, как нужно излагать свои мысли, что следует включить, а чем можно пренебречь, — которое развивается, только если постоянно иметь дело с наглядными примерами. Этого не достаточно для успеха, но это может пригодиться. Конечно, нет смысла копировать все без разбора — даже в самых лучших журналах можно найти массу неудачных приме-

ров. Разумнее в поисках примеров, на которых можно учиться, руководствоваться некими стандартами.

Вторая рекомендация — стараться писать простым языком. На основе сказанного выше могло сложиться неверное представление, что научная литература — это нечто недоступное для непосвященных, резко отличающееся от всех других видов литературы. Что читать научные работы *трудно*, поскольку они напичканы непонятными терминами, длинными, запутанными фразами, заумными доказательствами и т. д. Действительно, написание научного произведения требует следования определенного рода формальностям изложения, в чем, как правило, нет необходимости при написании произведений другого типа. Действительно, в любой науке есть свои специальные термины, употребление которых бывает предпочтительнее употребления менее точных слов бытового языка. Однако это не означает, что при написании научного произведения нужно *стремиться* к трудности; справедливо скорее обратное. В излагаемом материале будет масса трудностей; цель же автора — помочь читателю преодолеть эти трудности и понять содержание. Поэтому, в целом, короткие, простые слова предпочтительнее длинных и сложных; общеупотребимые слова предпочтительнее специальных терминов; а короткие и простые предложения предпочтительнее длинных и витиеватых. В конечном счете цель — сообщить информацию, а не поразить читателя своей искусственностью.

Третья рекомендация — стремиться к разнообразию. Простые предложения может быть и предпочтительнее, однако постоянное их использование быстро надоедает читателю. Более короткое слово — не всегда самое лучшее, и в любом случае, равномерное включение длинных слов и длинных предложений придает изложению своего рода ритм, улучшающий восприятие текста. Автор, постоянно использующий один и тот же набор слов, одну и ту же структуру предложения и абзаца, может быть, и преуспеш в ясности, но вряд ли заинтересует читателя. Ясность выражения мыслей и изящество стиля не исключают друг друга; нужно стремиться и к первому, и ко второму.

Четвертая рекомендация — стараться быть лаконичными. Запас времени и терпения читателей ограничен, а журналы жестко ограничены своим объемом. Статьи, длина которых непропорциональна их содержанию, вызывают негативную оценку. Представленные на рассмотрение в какой-нибудь из журналов, они тут же будут отвергнуты; в лучшем случае редакторы потребуют значительного сокращения текста. В процессе работы авторы должны постоянно задавать себе два вопроса. Необходимо ли включать данную информацию? Изложил ли я ее максимально эффективным способом? Или, как указывает Стернберг (Sternberg, 1993), автор должен быть всегда готов «увидеть различные варианты повторов!» (р. 58).

Совет быть лаконичными не означает, что научная работа должна быть написана телеграфным стилем. Краткость не имеет смысла, если теряется важная часть содержания или если немногословность автора мешает восприятию текста. Главной задачей остается донесение информации, а не экономия места. Разумнее руководствоваться следующим правилом: если есть сомнения, лучше включить. И автору, и читателю, как правило, легче отсеять лишнее в длинном пассаже, чем пытаться додумать короткое и непонятное предложение.

Последняя рекомендация — быть аккуратными. Каждый из нас время от времени пишет хуже, чем мог бы написать, просто потому, что не останавливается, чтобы хорошенько подумать, прежде чем написать или перечитать написанное. На

элементарном уровне, подавать работу на рассмотрение без *вычитки* совершенно непростительно. Статья, изобилующая опечатками, не только свидетельствует о неуважительном отношении к читателю, но и дает повод для негативных отзывов. На более высоком уровне важно перечитать написанное, с тем чтобы проверить логику изложения, грамматику и орфографию. Многие статьи и книги содержат явно ошибочные утверждения, противоречия, несоответствия и т. д., которые, скорее всего, были бы исключены, если бы автор нашел время перечитать свою рукопись. Не всякий может стать выдающимся стилистом, однако всякий может быть аккуратным и таким образом избежать при написании грубых оплошностей.

Разделы статьи

В этой части главы мы поговорим об основных составляющих психологического исследовательского отчета. Этот раздел не претендует на полный анализ всех вопросов, которые могут возникнуть; его задача — освещение лишь важнейших из них. Все рассмотренные здесь моменты более подробно изложены в *Manual APA*.

Введение

Главная функция «Введения» — сориентировать читателя в изучаемой теме. Характер ориентировки в некоторой степени зависит от сути исследования, поэтому требования того, что и в каком месте следует писать, не нужно понимать слишком буквально. Однако в целом «Введение» должно отвечать на несколько вопросов. В чем состоит изучаемая проблема? Почему она интересна для исследования? Каковы результаты предыдущих исследований по этой проблематике? Каковы недостатки предыдущих исследований, сделавшие необходимым дальнейшее изучение? На получение какой именно информации было направлено данное исследование? И как (в общих чертах) эта информация была получена?

Как правило, во «Введении» изложение идет в направлении от общего к частному. Поэтому «Введение» можно начать с определения области исследования (например, пиажетианское исследование процесса обучения), а затем переходить к специфическим вопросам, изучаемым в новом исследовании (например, можно ли обучать детей дошкольного возраста с помощью инструментального обусловливания?). Первая часть «Введения» обычно бывает посвящена обзору предшествующих исследований по данной тематике, заключительная часть чаще всего имеет непосредственное отношение к новому исследованию. Однако, как отмечалось, последовательность изложения не определена какими-то жесткими правилами; главное — ответить на вопросы, указанные в предыдущем абзаце, и сделать это ясно, доступно и убедительно.

При обзоре предшествующих исследований ключевым является слово «уместно». В исследовательский отчет нет необходимости включать полный литературный обзор. И дело не только в том, что на него не хватит места; вполне можно предположить, что читатели, которым адресован этот отчет, уже знакомы с общей историей вопроса. В случае с пиажетианским исследованием процесса обучения, к примеру, нет нужды распространяться о феномене сохранения или обосновывать его валидность; вполне можно рассматривать как данность то, что об этом известно любому, кого интересует подобный отчет. Нет необходимости и в том, чтобы стремиться осветить все предыдущие попытки развить представление о сохране-

нии; на сегодняшний день таких работ слишком много, и большинство из них будут лишь косвенно связаны с любым новым исследованием. Что имеет смысл, так это, если есть возможность, процитировать одну или несколько хороших обзорных статей, подводящих итог результатам, полученным в предыдущих работах. А что абсолютно необходимо, это рассмотреть все предшествующие исследования, которые непосредственно связаны с вашим собственным. Так, в примере с исследованием обучения нужно было бы упомянуть все работы с инструментальным обучением; то же касается и исследований с детьми дошкольного возраста. Оценить потенциальный вклад исследования можно только в контексте полного и тщательного рассмотрения того, что уже было сделано.

Заключительная часть «Введения», как правило, служит связующим звеном между предпосылками, с изложения которых начинается статья, и новым исследованием, которому посвящены все оставшиеся страницы. Именно здесь нужно разъяснить читателю все важнейшие из новых аспектов исследования. Здесь также уместно сформулировать конкретные гипотезы, если исследование действительно проводится в форме проверки гипотез. Даже если автор не хочет выдвигать никаких четких гипотез, бывает полезно поставить главные вопросы (например, можно ли обучить 3-летних детей?), на которые должно ответить исследование. Наконец, «Введение» чаще всего завершает краткое описание плана и процедур, которые были использованы при поиске ответов на поставленные вопросы. Подробности процедуры следует отложить до рубрики «Метод», и поэтому здесь их лучше опустить; однако введение в курс дела все-таки вполне уместно.

Метод

В разделе «Метод» рассказывается о том, что было проделано в исследовании. Эта часть обычно делится на несколько подразделов, по крайней мере на подразделы «Испытуемые» и «Процедура». В зависимости от характера исследования можно также включить отдельные разделы «Материалы», «Аппаратура», «План», «Подсчет результатов». В противном случае вся информация по этим вопросам объединяется разделом «Процедура». Выбор количества подразделов определяется сложностью исследования (например, насколько сложна использовавшаяся аппаратура) и предпочтениями исследователя.

Я уже говорил о той информации, которую следует представить в рубрике «Испытуемые»¹. Вспомним, что сюда должны входить не только демографические данные о выборке (например, количество, возраст, пол), но и подробности о методах отбора, процент согласившихся на участие из тех, к кому обращался исследователь, и процент не выбывших из исследования. В оставшейся части этого раздела основное внимание уделяется другим вопросам, которые должны рассматриваться в рубрике «Метод».

Часто высказывается мнение (в том числе и в *Manual APA*), что информация, включенная в «Метод», должна быть достаточно подробной, чтобы позволить вос-

¹ Если вы читали статьи по психологии, изданные до 1995 года, то, вероятно, отметили, что в отчетах об исследованиях, проводимых с людьми, термин *subject* (*субъект*) недавно заменил термин *participant* (*участник*). Цель этого изменения (которое является одним из множества новшеств в терминологии, введенных в последней версии *Manual*) — подчеркнуть, что люди, принимающие участие в исследовании, активно влияют на ход исследования, а не представляют собой лишь пассивные объекты изучения.

произвести исследование. Справедливость этого утверждения зависит от того, что подразумевается под словом «воспроизвести». Если этот термин употребляется в смысле «провести точно такое же исследование во всех деталях», тогда вряд ли какой из отчетов удовлетворит этому требованию, за исключением, возможно, отчетов о простейших исследованиях; на описание мельчайших подробностей процедуры просто не хватит места. Если же этот термин употребляется в смысле «проведения исследования того же *типа*, с сохранением всех *важнейших* аспектов процедуры», задача воспроизводимости становится вполне осуществимой. Трудность в том, чтобы определить, какие из аспектов процедуры важны и, следовательно, должны быть описаны в «Метод».

Есть множество оснований, по которым можно определить, что некая особенность процедуры *не* достаточно важна, чтобы включать ее описание в отчет. В любом исследовании есть ряд процедурных мелочей (например, сидел испытуемый слева или справа от экспериментатора), вероятность влияния которых на результаты столь мала, что их безбоязненно можно опустить. В исследованиях со словесными инструкциями редко возникает необходимость их дословного воспроизведения; как правило, достаточно передать общий смысл. В такого же рода исследованиях с множеством заданий, как правило, бывает достаточно в качестве примера описать одно или два из них, по которым можно судить об остальных. И так же, как и в случае с «Введением», можно рассчитывать на то, что определенная информация доступна любому читателю-специалисту. Если, к примеру, была использована стандартная аппаратура, достаточно будет указать ее название и, может быть, номер модели; читатели, которым данное оснащение незнакомо, легко смогут его найти. Аналогично, распространенные тесты не нуждаются в детальном описании; вполне хватит указания литературного источника и, возможно, краткого изложения сути.

Важнейшие детали «Метода» — это те, о которых читателю необходимо знать, чтобы понять и оценить исследование. В эту категорию попадает и упомянутая выше информация об испытуемых, а также ряд аспектов экспериментальной процедуры, аспектов, которые в этой книге рассматриваются, в качестве важнейших компонентов исследования. По какому принципу испытуемых распределяли по экспериментальным группам? В чем конкретно состояли экспериментальные условия и в каком виде они представлялись испытуемым? (Заметьте, что здесь дословное описание *может* оказаться уместным.) Что выступало в качестве зависимых переменных и как они измерялись? Каков был порядок экспериментальных событий? Использовался ли «слепой» метод? Сколько было задействовано наблюдателей или тестирующих? Оценивалась ли надежность, и если да, то каков был ее уровень? Конкретные сведения, которые необходимо предоставить, в известной мере зависят от характера исследования. Определяя то, о чем нужно упомянуть, полезно ориентироваться на критерии воспроизводимости. Однако действительно решающий критерий — это *оценка*: достаточно ли будет информации читателю, чтобы оценить валидность метода?

Результаты

Раздел «Результаты» содержит сведения о том, что было обнаружено в ходе исследования. В нем приводятся значения как дескриптивных статистических показателей — к примеру, средний уровень выполнения заданий в разных условиях, — так

и статистические показатели, выводимые логическим путем, — например, критерии значимости при сравнении средних значений. Необходимо включить оба вида статистических показателей. Как мы узнали из главы 7, дескриптивные показатели довольно трудно интерпретировать при отсутствии соответствующих статистических проверок. Аналогично F -критерий или уровень p имеют ограниченное значение при отсутствии информации о средних, на которых они основаны. Читатель должен знать и формы статистического анализа, и данные, которые использовались при анализе.

Так же как и с «Введением» и «Методом», исследователь должен решить, какую информацию включить в «Результаты» необходимо, а какую можно опустить. Если предыдущие разделы выполнили свои функции, возникнет понимание того, какие именно вопросы будут освещены в «Результатах». На каждый из этих вопросов должен быть получен какой-то ответ, даже в тех случаях, когда ответ состоит из признания, что никаких значимых эффектов не обнаружено. Кроме того, нет такого правила, согласно которому обо всех формах анализа, проведенных в разделе «Результаты» должно быть известно заранее; бывает, что появляются неожиданные или случайные результаты, и если это происходит, их необходимо включить в отчет. С другой стороны, нет и такого правила, согласно которому необходимо указать результаты *всех* проведенных аналитических процедур. Исследователь должен избегать заполнять свое отчет второстепенными результатами, ничего не прибавляющими к выводам исследования.

Раздел «Результаты» — наиболее трудная для чтения часть статьи. Автор может предпринять некоторые шаги, чтобы облегчить задачу читателя. Один из способов — изложение в логической последовательности. Как правило, начинают с результатов главных аналитических процедур и только затем указывают менее существенные результаты. Если во «Введении» формулировались некие гипотезы, результаты следует излагать, ссылаясь на эти гипотезы, и по очереди указывать результаты проверки каждой из них. Если производилась оценка нескольких зависимых переменных, имеет смысл представить результаты для каждой из зависимых переменных отдельно. Если использовались разные статистические процедуры (например, дисперсионный и корреляционный анализ), результаты следует структурировать в соответствии с типом анализа. Опять-таки выбор оптимального подхода зависит от конкретного исследования. Важна логичность изложения и уверенность в том, что и читателю ясна эта логика.

Автору также следует сделать некоторые обобщения, чтобы помочь читателю лучше разобраться в данных. Существует заблуждение, что раздел «Результаты» должен состоять из сухих цифр и значений статистических критериев. Действительно, функция раздела «Результаты» — представление данных, тогда как функция раздела «Обсуждение» — их интерпретация. Однако из этого *не* следует, что в разделе «Результаты» нельзя говорить о значении данных. Наоборот, разобраться в непрерывной последовательности t , F , коэффициентов корреляции довольно трудно и скучно. Поэтому лучше всего вставить одно-два пояснения — либо до представления результатов анализа, либо в качестве краткого итога после — и на простом английском рассказать, что означают эти цифры.

Важно ясно представлять себе то, о чем здесь идет речь. В «Результатах» допустима своего рода первичная интерпретация, не отступающая далеко от данных. Вполне уместно, например, сказать: «Было выявлено значимое взаимодействие между фактором возраста и экспериментальными условиями, $F(1,96) = 8,95$, $p < 0,01$ (соответствующие средние см. в табл. 9.1). Обучение действительно произвело значимый эффект, но только в группе 5-летних детей; у 3-летних детей не наблюдалось никакого прогресса». Однако совершенно *неуместно* было бы продолжить: «Большая восприимчивость старших детей согласуется с точкой зрения Пиаже, подчеркивавшего значение операциональных структур...» Интерпретации более высокого порядка нужно приберечь до раздела «Обсуждение».

Автор может помочь своим читателям и при помощи рационального использования таблиц и схем. Здесь важно слово «рациональное»: печатание таблиц и схем стоит довольно дорого, и поэтому редакторы неодобрительно относятся к злоупотреблению этими вспомогательными средствами. Оправданно или нет использование таблицы или схемы, зависит от того, насколько легко можно донести необходимую информацию в тексте. Если должно быть представлено лишь небольшое количество средних, в таблице нет нужды, поскольку несколько чисел легко можно включить в текст. Однако при наличии 15 или 20 средних читателю значительно удобнее воспринять данные в виде таблицы, а не последовательности цифр в тексте. В целом, таблицы предпочтительнее использовать для представления точных значений и оперирования большим объемом данных (как в случае с 15 или 20 средними). Схемы особенно удобны для отображения трендов и взаимодействий.

При использовании таблиц и схем следует руководствоваться еще несколькими правилами. Таблицы и схемы не должны дублировать данные, представленные в тексте или в других таблицах или схемах. Если одна и та же информация появляется дважды, один из источников следует исключить. Таблица и схема должны объяснять сами себя — то есть должны быть понятны без ссылок к подробному описанию в тексте. Наконец, хотя таблицы и схемы должны объяснять сами себя, тем не менее необходимо указать, о чем свидетельствуют представленные в них данные. Таблицы и схемы дополняют текст, но не заменяют его.

Обсуждение

Раздел «Обсуждение» дает возможность подвести итоги и сформулировать выводы, к которым, по мнению автора, должны были прийти читатели. Этот раздел тесно связан и с «Введением», и с «Результатами». Во «Введении» намечаются вопросы, на изучение которых направлено исследование; в «Обсуждении» обобщаются ответы, которые, с точки зрения исследователя, дало исследование. На что бы ни надеялся, что бы ни ожидал получить автор, решающей детерминантой содержания «Обсуждения» являются данные.

Для многих авторов «Обсуждение» — эта часть статьи, которую труднее всего писать (особенно когда результаты оказываются совсем не такими, как ожидалось!). Ниже я дам несколько рекомендаций, чего следует всячески избегать при написании раздела «Обсуждение» (и отслеживать при прочтении).

В обсуждении не стоит еще раз перечислять показатели, представленные в разделе «Результаты». Перед тем как интерпретировать результаты, вполне уместно

их несколько обобщить. Однако предоставить информацию о данных — это задача раздела «Результаты». Задача же «Обсуждения» — интерпретация этих данных.

За очень редким исключением предварительно не указанные результаты не должны анализироваться в «Обсуждении». Все новые данные должны вводиться в разделе «Результаты».

Как и в других разделах статьи, изложение в «Обсуждении» должно быть ясным и логичным. Некоторые студенты пишут «Обсуждение» в так называемой манере «стриптиза», начиная с второстепенных деталей, постепенно подходя к главным выводам. Обычно для читателей удобнее, когда главные выводы сформулированы в начале «Обсуждения». Уточнения, ограничения и тонкие моменты можно проработать после.

В «Обсуждении» нет необходимости уделять равное внимание всем гипотезам или всем результатам. Одна из обязанностей автора — выделить главное. Однако слишком большая избирательность может привести к искаженному представлению результатов исследования. Неправильно, к примеру, все внимание уделять одному результату, который подтвердил определенную гипотезу, в то же время игнорируя три другие результата, которые ее не подтвердили.

В разделе «Обсуждение» вполне уместно несколько абстрагироваться от непосредственно имеющихся данных и позволить себе сделать ряд далеко идущих выводов. Однако некоторые авторы позволяют себе абстрагироваться настолько, что теряют из вида само исследование, которое они призваны обсудить. К таким витиеватым рассуждениям иногда прибегают при попытке компенсировать недостаточную убедительность результатов. В любом случае, все замечания, сделанные в «Обсуждении», должны иметь ясную связь с исследованием, которому посвящен отчет.

Обычно «Обсуждение» завершают рассмотрением перспектив дальнейшего исследования. Это отличный способ, если предложения автора: а) конкретны, б) разумны и в) изложены в сжатой форме. Однако просто сказать, что требуется дальнейшее изучение — значит ничего не сказать.

Еще один широко используемый метод завершения «Обсуждения» — рассмотрение недостатков проведенного исследования, которые, по мнению автора, отразились на результатах. Это также отличный способ; указание на слабые места можно считать проявлением элементарной честности автора перед читателями. Однако не нужно излишне подробно останавливаться на недостатках, защищаться или оправдываться. Поскольку, если «Обсуждение» посвящено в основном недостаткам, для читателя естественно задаться вопросом: «Зачем вообще серьезно относиться к этой работе?» А естественный вопрос, который возникнет у редактора: «Зачем вообще публиковать этот отчет?»

Другие разделы

Хотя «Введение», «Метод», «Результаты» и «Обсуждение» составляют основу исследовательского отчета, перечень его значимых частей ими не исчерпывается. В табл. 9.1 представлен более полный список разделов исследовательского отчета. Эти разделы указаны в том порядке, в котором они должны быть представлены на рассмотрение в редакцию журнала. Опять-таки более подробно о содержании каждого из разделов можно узнать из *Manual APA*.

Таблица 9.1

Разделы исследовательского отчета

Раздел	Описание
Титульный лист	Название должно быть кратким (10–12 слов), ясным и информативным. Избегайте лишних слов, таких как «Исследование, направленное на изучение...» Титульный лист должен также включать сведения об авторе, информацию о специальности и месте работы автора, а также его подпись
Краткий обзор	Описание исследования в 100–150 слов (в зависимости от журнала). Должен включать информацию об изучаемой проблематике, методе исследования, основных результатах и главных выводах
Введение	Во «Введении» автор знакомит читателей с изучаемой проблемой, обосновывает интерес к этой проблеме и дает оценку современному состоянию знаний по этой теме, основанную на обзоре соответствующих исследований, проведенных ранее. Раздел завершает беглое рассмотрение предлагаемого исследования, включая конкретные цели и конкретные гипотезы
Метод	Краткое описание процедуры проведения исследования, включая информацию об испытуемых, аппаратуре и материалах, а также об экспериментальных манипуляциях. Описание должно быть достаточно подробным, чтобы дать возможность оценить исследование
Результаты	В этом разделе представляются главные из полученных данных, а также результаты статистического анализа этих данных
Обсуждение	Интерпретация результатов исследования. Помещение данных в теоретический контекст, представленный во «Введении». В «Обсуждении» следует попытаться ответить на конкретные вопросы и оценить конкретные гипотезы, сформулированные во «Введении»
Список литературы	Все источники, на которые делалась ссылка в тексте, представленные в алфавитном порядке
Приложение	В «Приложении» дается достаточно объемный материал (например, математические доказательства, описание стимульного материала), который не подходит для представления в тексте. Использование «Приложения» по возможности следует избегать
Авторское примечание	Информация об авторах, а также выражение признательности лицам и организациям, оказывавшим поддержку
Сноски	Материалы относительно небольшого объема, которые не подходят для включения в текст. Использование сносок по возможности следует избегать
Таблицы	Объединяют данные, которые удобнее поместить в табличной форме, а не в тексте
Пояснения к схемам	Поясняют схемы, включенные в статью
Схемы	Объединяют данные, которые удобнее всего представить в виде схемы

Частные аспекты стилистики

Нормы АПА

Как уже отмечалось, в этом разделе мы будем избирательны и уделим внимание лишь нескольким правилам, которые студенты нарушают в своих работах чаще всего.

Заголовки. Стил ь АПА предусматривает пять уровней заголовков. Эти уровни иллюстрирует табл. 9.2. Если необходим только один заголовок (обычно в очень коротких статьях), используется уровень 1, помещенный в центр заголовка прописными и строчными буквами. Если необходимы два заголовка, используются уровни 1 и 3. Если необходимы три уровня, добавляется заголовок абзаца, помещенный с отступом. Четыре или пять заголовков используется редко, в основном в отчетах о широкомасштабных исследованиях, в длинных обзорах и в статьях теоретического плана.

Таблица 9.2

Уровни заголовков в журналах АПА

Тип заголовка	Пример
ПОМЕЩЕННЫЙ В ЦЕНТРЕ ЗАГЛОВОК ПРОПИСНЫМИ БУКВАМИ (УРОВЕНЬ 5)	ЭКСПЕРИМЕНТ 1: ИССЛЕДОВАНИЕ ВАЛИДИЗАЦИИ МЕТОДА ИНТЕРВЬЮ
Помещенный в центре заголовка прописными и строчными буквами (уровень 1)	Внешняя валидизация
Помещенный в центре, подчеркнутый заголовок прописными и строчными буквами (уровень 2)	Метод
Помещенный вплотную к левому краю подчеркнутый боковой заголовок прописными и строчными буквами (уровень 3)	Испытуемые
Помещенный с отступом, подчеркнутый заголовок абзаца строчными буквами, оканчивающийся точкой	Неклиническая группа.

Измерения. Требование АПА — выражать результаты измерения в метрических единицах, то есть в миллиметрах, а не в дюймах, в метрах, а не в ярдах и т. д. Результаты измерений, представленные в неметрических единицах, должны сопровождаться переводом в метрические эквиваленты — к примеру: «Длина палки составляла 3 фута (0,91 м)».

Числа в тексте. Согласно правилам АПА, числа от нуля до девяти обычно пишутся словами, а двузначные и многозначные числа — цифровыми символами. К примеру, «Было восемь экспериментаторов и 120 испытуемых». Однако из этого правила есть и ряд исключений. Наиболее характерное исключение состоит в том, что единицы измерения должны сопровождать цифровые символы, а не слова (то есть, «4-летний возраст», «6 недель» и т. д.), а предложения нельзя начинать с

цифрового символа (то есть «Сто двадцать испытуемых были включены в экспериментальную группу»). Другие исключения подробно описаны в «Руководстве АПА». По своему опыту скажу, что ошибки с числами чуть ли не самые частые ошибки в работах студентов; поэтому, используя числа, особенно важно сверяться с «Руководством».

Описание в тексте статистических процедур. Автор должен всегда указывать на то, какие процедуры он использовал для анализа каждого из аспектов данных. Например, «Правильные ответы на посттест анализировались с помощью дисперсионного анализа, где в качестве независимых переменных выступали возраст и экспериментальные условия». Приводить научные источники для статистических процедур нет необходимости, за исключением тех случаев, когда процедура довольно оригинальна или ее использование может вызвать критические замечания. Результаты проведения логического анализа должны включать название статистического показателя, степени свободы, значение статистического показателя и уровень вероятности — например, «Был обнаружен значимый эффект возраста, $F(1,96) = 7,90, p < 0,01$ ». Заметьте, что результаты логического анализа выделяются запятыми, а не скобками. Соответствующие дескриптивные статистические показатели должны включать меры центральной тенденции (обычно среднее), а также какую-либо из мер изменчивости (обычно стандартное отклонение) — например, «Старшие дети ($M = 6,88, SD = 2,45$) справились с заданием лучше, чем младшие ($M = 4,90, SD = 2,26$)»¹.

Ссылки в тексте. Форма, принятая АПА для обозначения в тексте ссылок, используется и на страницах этой книги: фамилия автора и года публикации, заключенные в скобки, — например, «(Smith, 1992)». Если фамилия автора является частью текста, в скобках указывается только год издания — например, «Смит (1992) сообщает, что...». Если книга создана двумя соавторами, всегда указывается обе фамилии; в тексте их соединяет «и», а в скобках — знак «&». Если работа имеет от трех до пяти соавторов, при первом упоминании приводятся все фамилии, а при последующих указывается фамилия только одного из соавторов, за которой следует «и др.» («et al.»). Если у работы шесть и более соавторов, «и др.» используется уже при первом упоминании. Если в скобках указывается фамилия одного и того же автора, ссылки выстраиваются в хронологическом порядке; если в скобках заключены фамилии *разных* авторов, ссылки выстраиваются в *алфавитном* порядке. Например, «(Smith, 1984, 1989, 1992)» и «(Brown, 1979; Jones, 1974; Smith, 1992)».

Список литературы. Ссылка на каждый из источников, упоминавшихся в тексте, выносится в конец статьи, в раздел «Список литературы». В него включают *только* те источники, которые упоминались в тексте. Автор должен убедиться в точности информации, представленной в «Списке литературы», чтобы читатели, заинтересовавшиеся тем или иным источником, могли легко его найти. Кроме того, автор должен следовать правилам АПА для «Списка литературы», которые занимают целые 60 страниц *Manual*!

¹ В этом разделе все статистические обозначения выделены курсивом, как это делается при публикации научной статьи в журнале. Заметьте, однако, что в рукописи, предоставляемой на рассмотрение в редакцию журнала, следует использовать подчеркивание, а не курсив. Это касается не только статистических обозначений, но и всех случаев, когда в опубликованной версии должен появиться курсив.

Общие вопросы английского языка

Изложенный ниже материал носит даже еще более избирательный (и своеобразный) характер. Здесь выборочно представлены правила английского языка, которые студенты чаще всего нарушают в своих письменных работах (и не только в них).

Affect—Effect. Оба слова, и *affect* (воздействовать; аффект), и *effect* (производить; эффект, результат) могут употребляться и как существительные, и как глаголы. Однако чаще всего *affect* имеет значение глагола «влиять», а *effect* — значение существительного «исход» или «результат». Таким образом, следует, к примеру, говорить «Did the manipulation affect performance? The posttest showed a clear effect».

Сокращения с апострофом. Не используйте сокращения.

Data (данные). *Data* (данные) — существительное множественного числа. Поэтому нужно говорить «The data were clear» («Данные были однозначны»), а не «The data was clear». Аналогично, *criteria* (критерии) и *phenomena* (феномены) — существительные множественного числа, и поэтому согласуются с глаголами во множественном числе. Формы единственного числа — *datum* (используется редко), *criterion* и *phenomenon*.

e. g.—i. e. Сокращение *e. g.* означает «например», а сокращение *i. e.* — «то есть». Поэтому правильно писать: «Многие теоретики (*e. g.* Brown, 1979; Smith, 1992) заявляют, что...» и «Решающее значение имели экспериментальные манипуляции, то есть (*i. e.*) наличие или отсутствие награды».

Fewer—less. *Fewer* (меньше) употребляется в отношении исчисляемых объектов, а *less* (менее) характеризует неисчисляемые понятия. Поэтому нужно говорить «The control participants gave fewer correct responses» («Испытуемые из контрольной группы дали меньше правильных ответов») и «The control participants showed less success» («Испытуемые из контрольной группы менее успешно справились с заданием»). Очень частая ошибка — использование *less* вместо *fewer*, например, «The control participants gave less correct responses» («Испытуемые из контрольной группы дали менее правильных ответов»).

Дефис. Иногда бывает трудно решить, разделять составное слово дефисом или нет, и в *Manual APA* указаниям по этому вопросу посвящено 5,5 страниц. В соответствии с нормами стиля АПА, некоторые префиксы пишутся без дефиса; в эту категорию попадают *inter*, *mid*, *multi*, *non*, *post*, *pre* и *semi*.

It's—its. *It's* — это сокращенная форма «it is» или «it has»; *its* — это притяжательное местоимение, означающее «принадлежащее этому». Ошибочно, будучи уверенными в том, что притяжательная форма подразумевает наличие апострофа, использовать в качестве притяжательного местоимения *it's*.

Only (только). Хотя существуют и исключения, обычно слово *only* («только») употребляется непосредственно перед словом, которое оно характеризует, поэтому нужно говорить: «The failure manipulation produced visible distress in only two children» («Создание ситуации неизбежности неудачи вызвало заметный дистресс только у двух детей»), а не «The failure manipulation only produced visible distress in two children» («Создание ситуации неизбежности неудачи только вызвало заметный дистресс у двух детей»).

Местоимения. С употреблением местоимений могут возникнуть разнообразные проблемы. Согласно «Руководству АПА», сегодня в научных работах использование местоимения первого лица единственного числа вполне приемлемо и даже предпочтительнее использования исключительно третьего лица и страдательного залога. Однако студенты должны отдавать себе отчет, что местоимение «я», хотя и используется сегодня чаще, чем раньше, все же не столь распространенная форма; к ней прибегают в основном заслуженные (или весьма уверенные в себе) исследователи, а также авторы теоретических работ. Использование «я» в банальном исследовательском отчете может вызвать у читателей раздражение. Использование же «мы» в коллективных исследовательских отчетах — вполне типичное явление (заметьте, однако, если у работы только один автор, употреблять авторское «мы» не следует). Интересную точку зрения на вопрос первое/третье лицо содержится в работе Polysen, Levinson, & Miller, 1982.

В *Manual* также высказывается пожелание избегать использования сексистских терминов, включая и сексистские местоимения; в действительности, одно из новшеств последней версии — более подробное рассмотрение проблемы не только сексистских выражений, но и предвзятости в словах в целом. Студентам следует внимательно ознакомиться с этим разделом и обращаться к нему при написании исследовательских работ. Важно, однако, чтобы стремление к несексистской речи не отражалось на грамматической правильности или ясности изложения. Поэтому следует избегать предложений типа «Ребенку велели, чтобы он/она нажал(а) на кнопку, как только увидит на расployенном перед ним/ней экране вспышку». Следует также избегать произвольного изменения рода местоимений. Поэтому, если в начале статьи пол ребенка был указан как женский, в дальнейшем должно использоваться местоимение «она» (хотя в более длинных работах, таких как эта книга, в качестве альтернативы использования местоимения только одного рода или иногда неуклюжего «он/она» допускаются переходы). Наконец, при обозначении ребенка, даже младенца, никогда не должно использоваться местоимение «it» («оно»).

Последнее, что можно сказать о местоимениях, имеет отношение к элементарной грамотности: местоимение должно согласовываться в числе с антецедентом: поэтому нужно говорить: «Каждый испытуемый заполнил свой бланк для ответов», а не «Каждый испытуемый заполнил свои бланки для ответов» («Each participant filled out his or her response sheet», а не «Each participant filled out their response sheet»).

Время. В журнальных статьях сообщается о событиях, которые уже произошли. Поэтому в большинстве случаев уместнее всего употреблять прошедшее время — то есть «Смит представил данные...», «Испытуемым сказали...», «Анализ показал...» Настоящее время можно использовать при обсуждении результатов, которые представлены в тексте, например: «Из таблицы явствует...» Настоящее время употребляется также в утверждениях, которые по сути своей не связаны со временем — «Теория гласит», «Маленькие дети часто бывают импульсивными», «Понимание принципа сохранения является одной из важнейших способностей».

Резюме

Последний этап исследовательского проекта — сообщение результатов проделанной работы другим людям. В этой главе рассматривается вопрос составления исследовательского отчета.

Глава начинается с того, что подчеркивается необходимость ответственного отношения к написанию научной работы; отмечается, что ясность и убедительность изложения во многом определяют значение, которое будет иметь научная работа. Хорошо написанная статья имеет больше шансов на опубликование, на внимательное отношение читателей и научной общественности. Предлагается ряд рекомендаций по повышению эффективности составления письменной работы. Первая рекомендация — учиться тому, о чем и как говорить, на примерах грамотных психологических работ. Вторая рекомендация — стремиться к простоте, поскольку злоупотребление специальными терминами и неоправданная сложность мешают ясному выражению мыслей. Третья рекомендация — стараться разнообразить свою работу, стремясь сделать материал не просто понятным, но и интересным. Четвертая рекомендация — стремиться к лаконичности, поскольку излишняя детализация и многословие могут быстро исчерпать и отведенный журналом объем, и терпение читателя. Последняя рекомендация — будьте аккуратны, так как если найти время и удостовериться в том, что все части рукописи соответствуют замыслу, то удастся избежать многих грубых оплошностей.

В средней части главы рассматриваются основные разделы психологического исследовательского отчета. При описании каждого раздела затрагивается два вопроса: какой материал следует включить, и в какой форме его преподнести. Во «Введении» необходимо представить проблемы и цели, лежащие в основе исследования; поместить исследование в контекст уже проделанной в этом направлении работы и подготовить читателей к знакомству с предлагаемым исследованием. В разделе «Метод» описываются испытуемые и процедуры, использовавшиеся для сбора данных. Это описание должно быть достаточно подробным, позволяющим дать исследованию критическую оценку. В разделе «Результаты» представляются данные, полученные в ходе исследования; этот раздел включает в себя как дескриптивные статистические показатели (например, средний уровень успешности выполнения задания), так и статистические показатели, выводимые логическим путем, (например, значение критериев значимости различий между средними), а также краткие выводы о том, что означают эти показатели. В «Обсуждении» автор интерпретирует результаты исследования, дает ответы на вопросы, поставленные «Введением», оценивает все выдвинутые гипотезы и подводит итог тому, что удалось, а чего не удалось достичь в исследовании.

Заключительный раздел главы посвящен частным вопросам стиля, которые часто вызывают у студентов затруднения. Рассматриваются как аспекты стиля АПА (например, правила для заголовков и списка литературы), так и аспекты английской грамматики — «i. e., don't write passages like this one, which would be likely, even if they contained less errors (and there are five here), to effect the reader negatively» («т. е. не пишите таких фраз, как эта, что, даже если бы в них было менее ошибок (а в этом их пять), могло бы иметь негативный эффект на читателя»).

Упражнения

С согласия преподавателя обменяйтесь письменными работами с одним из своих однокурсников. Подготовьте подробный критический отчет о работах друг друга, оценивая содержание, стиль и соблюдение норм АПА.

В главе 2 в качестве примера приведено два исследования, которые в дальнейшем не раз упоминаются в тексте: исследование памяти и времени запоминания у младших школьников, проведенное Дюфресном и Кобасигавой (Dufresne & Kobasigawa, 1989), а также исследование пространственной памяти у молодых и пожилых лиц, проведенное Парк и Черри (Cherry & Park, 1993). Используя описания, приведенные в тексте, а также собственную изобретательность (поскольку вам придется додумать те подробности, о которых не было ничего сказано), составьте раздел «Метод» для одного из этих исследований и раздел «Результаты» — для другого. Сравните свою работу с соответствующими разделами первоисточников.

Младенчество

В последних четырех главах мы перейдем от общих принципов исследования к специфическим вопросам психологии развития. Структура дальнейшего изложения носит частично проблематический, а частично хронологический характер. Эта часть книги начинается с главы, посвященной младенчеству, и заканчивается главой, посвященной пожилому возрасту. Между ними — главы, в которых речь идет о когнитивном и социальном развитии. Хотя такое проблематически-хронологическое изложение может показаться нелогичным, оно в действительности отражает ситуацию, сложившуюся в данной области исследования. Одни психологи определяют свою специальность преимущественно с точки зрения изучаемых ими возрастных групп; другие — преимущественно с точки зрения разрабатываемой проблематики. Аналогично, одни методологические вопросы чаще встают при изучении разных возрастных групп, а другие — более тесно связаны с конкретной проблематикой.

Эта глава состоит из двух больших разделов. Начнем мы с рассмотрения некоторых общих вопросов проведения исследования с младенцами. По большей части суть этих вопросов сводится к применению на практике того, о чем говорилось в первых шести главах. Однако, как мы увидим, исследователь, решивший изучать младенцев, сталкивается с особыми методологическими трудностями.

Второй, и более длинный, раздел этой главы посвящен конкретным видам изучения периода младенчества. Освещение данной тематики, что, впрочем, справедливо по отношению ко всем оставшимся четырем главам, неизбежно будет избирательным. Тем не менее вполне можно обсудить важнейшие и в методологическом плане наиболее интересные исследования.

Общие вопросы

Отбор

Начнем с одного из основных этапов исследовательской работы, описанных в главе 1, — с привлечения испытуемых. В главе 1 отмечалось, что для исследователей в области психологии развития эта проблема приобретает особое значение. Добавим

теперь, что исследователь, занимающийся младенцами, обрекает себя на самые серьезные трудности при поиске испытуемых. Младенцы, в отличие от детей старшего возраста и взрослых, не могут изъявить желания принять участие в эксперименте. Нет готовой выборки младенцев в специальных учреждениях, таких как школа, класс изучения основ психологии или детский сад. Как же тогда исследователю найти малышей для изучения?

Ответ следующий — разными способами, при этом выбор способа зависит от возраста младенца и ресурсов исследователя. Если объект изучения — новорожденные, *можно* обратиться в специальные учреждения и обследовать малышей, пока их не забрали из роддома. Заметьте, однако, что в связи с устойчивой тенденцией к уменьшению времени пребывания новорожденных и их матерей в роддоме последнее время обследование в роддоме становится все более трудновыполнимой задачей. Если же исследователь намеревается изучить малышей постарше, для привлечения испытуемых можно использовать списки детей, поставленных на учет у педиатров и в профилактических детских учреждениях. В таких случаях список фамилий — это лишь отправная точка; исследователь должен еще получить согласие на обследование ребенка. Детей любого возраста можно найти, подав в газету объявление с просьбой откликнуться заинтересовавшихся родителей. Можно также действовать в противоположном направлении: некоторые специалисты отслеживают сообщения о рождении детей, а затем, по достижении ребенком интересующего исследователя возраста, звонят его родителям. Каков бы ни был подход, непереносимое условие успеха одно и то же — готовность родителей позволить обследовать своего малыша.

Как следует из вышесказанного, поиск младенцев-испытуемых может потребовать больших затрат времени и усилий. Однако действительная проблема — это репрезентативность полученной выборки. Я уже говорил в главе 2, что по-настоящему случайный отбор из изучаемой популяции — это труднодостижимый идеал. Некоторое отклонение от случайности — норма для исследования. Однако, по крайней мере при изучении детей, наиболее значительны такие отклонения при работе с младенцами. Малыши, которых приносят в детские профилактические учреждения, — это одна подгруппа популяции; малыши, которых приносят к педиатру, — другая. Родители, позитивно реагирующие на телефонные звонки с просьбой об участии в исследовании, — одна подгруппа; те, кто откликается на объявления в газетах, — другая, несомненно, не столь многочисленная подгруппа. В целом, есть основания полагать, что родители, соглашающиеся на участие в исследовании своих малышей, более образованны, чем родители в своем большинстве; иногда они также испытывают серьезные денежные затруднения и поэтому достаточно охотно соглашаются на предложения, связанные с материальным вознаграждением за участие (заметьте, что родители-студенты попадают в обе эти категории!). По причине всего вышесказанного малыши, в конце концов попадающие в исследование, составляют выборку, созданную с существенными нарушениями принципа случайного отбора.

Второй вопрос касается количества малышей, которые пройдут все исследование до конца. В исследованиях младенцев доля выбывающих весьма значительна. В действительности уровень выбывания в исследованиях младенцев чуть ли не самый высокий среди всех исследований, проводимых с людьми; такие показатели, как 50–60 % не являются чем-то из ряда вон выходящим. Чаще всего в качестве

причины выбывания называют нервность (что зачастую — всего лишь эфемеризм); сонливость и засыпание — также типичная проблема. Поэтому выборка, изначально созданная с отклонением от принципа случайного отбора, еще более отклоняется от этого критерия, настолько, насколько выбывшие из исследования малыши закономерно отличаются от оставшихся.

Является ли отклонение от критерия случайности в большинстве выборок младенцев серьезной проблемой? Необязательно. Тот факт, что выборка не отвечает критерию случайности, не означает, что полученные результаты нельзя перенести на некоторую большую группу. Об этом уже говорилось в главе 2 при обсуждении вопроса отбора в целом. Как отмечалось, при отборе важна не случайность сама по себе, а то, чтобы отклонение от случайности не искажало результаты. Во многих исследованиях вероятность того, что отклонение от случайности отразится на результатах, просто ничтожна. По отношению к младенчеству и к вопросам, изучаемым на выборке младенцев, это справедливо даже в большей мере, чем по отношению к довольно неоднородным группам старших детей. Это также справедливо по отношению не только к первичному отбору, но и к проблеме выбывания. Все малыши время от времени бывают сонливыми или нервными. Одних малышей эксперимент застает в тот момент, когда они нервны или сонливы, и поэтому они выбывают; других эксперимент застает в тот момент, когда они не спят и не капризничают, поэтому они остаются. Но нет никаких оснований полагать, что эти группы как-то закономерно различаются друг от друга.

Хотя довод о том, что отклонение от случайности не имеет значения, обычно справедлив, его применение к конкретным случаям все же требует тщательного изучения. В частности, вполне оправдано поставить под сомнение внешнюю валидность, когда выборка весьма своеобразна или уровень выбывания крайне высок. Кроме того, есть данные, позволяющие поставить под вопрос правомерность допущения, что выбывание в исследованиях младенцев носит случайный характер и поэтому не приводит к искажению результатов. Льюис и Джонсон (Lewis & Johnson, 1971) обследовали 3- и 6-месячных младенцев с помощью набора методик оценки зрительного внимания. Они сравнивали результаты двух групп малышей: 22 младенцев, прошедших всю серию тестов, и 15 младенцев, которые выполнили часть тестов, однако выбыли до окончания обследования. Младенцы, выполнившие всю серию тестов, демонстрировали большую способность различать простые и сложные стимулы; эта закономерность проявлялась как по параметру зрительных фиксаций, так и по параметру изменения частоты сердечных сокращений. На основе этих результатов Льюис и Джонсон (1971) сделали вывод о том, что «наши внутренние допущения о специфике группы младенцев, данные которых исключаются из анализа, вполне могут не соответствовать действительности... Выбывание младенцев, особенно в больших количествах, может привести к серьезному искажению полученных данных» (р. 1055). (См. также Richardson & McCluskey, 1983).

Значение состояния младенца

Младенец — это трудный испытуемый, который нередко бывает слишком беспокойным или слишком сонливым, чтобы с ним можно было довести эксперимент до конца. Это замечание имеет и более широкий смысл, поскольку указывает на значение, которое для понимания поведения младенца имеет его **состояние**. То, как малыш реагирует на окружающий мир, во многом зависит от его уровня активности. Нахо-

дится ли малыш в состоянии активного бодрствования? Или он вот-вот уснет? Голоден и поэтому раздражителен? Отчаянно рыдает? Или, возможно, сладко спит?

Как следует из вышеперечисленных вариантов, использовавшиеся до этого категории «нервный» и «сонный» обозначают состояние довольно приблизительно. Психологи в действительности разработали более тонкую градацию. Одна из распространенных классификаций представлена в табл. 10.1.

Таблица 10.1

Уровни активности младенца

Состояние	Описание
Глубокий сон	Во время глубокого сна малыш лежит спокойно с закрытыми глазами и неподвижными веками. Его дыхание ровное, а кожа — бледная
Поверхностный сон	Во время поверхностного сна мышечная реакция малыша на давление сильнее, чем во время глубокого сна. Малыш непроизвольно вздрагивает, ворочается; лицо подвижно. Кожа может быть розовой, дыхание неровное
Периодический сон	Периодический сон — это сочетание глубокого сна с поверхностным: он состоит из периодов частого дыхания и подергивания, сменяющихся периодами полного спокойствия
Состояние сонливости	В состоянии сонливости малыш умеренно активен. Он то закрывает, то открывает глаза; взгляд тусклый; дыхание ровное, но более частое, чем при глубоком сне
Пассивное бодрствование	Малыш не спит, но остается пассивным. Его глаза открыты, взгляд ясен; малыш с интересом смотрит по сторонам. Тело относительно расслаблено, а дыхание довольно частое и неровное
Активное бодрствование	В период активного бодрствования малышу трудно сосредоточиться, у него бывают вспышки бурной активности. Во время этих вспышек он двигает ножками и всем телом. Продолжительность и интенсивность таких вспышек активности бывает разной
Плач	Состояние плача характеризуется крайне высоким уровнем активности. Кожа малыша краснеет, и он плачет (хотя пока еще без слез)

Источник: *Understanding Infancy* (p. 38, 40), E. Willemsen, 1979, San Francisco: W. H. Freeman & Co.

Разумеется, состояние испытуемого имеет значение не только в выборе младенцев. Любой из нас более восприимчив к сигналам окружающей среды, будучи в состоянии ясного сознания, а не в глубоком сне; и любой из нас ведет себя несколько иначе, когда у него сводит от голода желудок. Однако по разным причинам при изучении младенцев состояние имеет большее значение, чем при изучении старших детей или взрослых. В младенчестве состояние более лабильно; переходы от одного состояния к другому происходят довольно часто. Как следует из данного факта, младенцы гораздо хуже контролируют свое состояние, чем старшие дети или взрослые. Можно не без оснований предположить, что 10-летний ребенок, находящийся в состоянии бодрствования в начале эксперимента, не уснет и через 20 ми-

нут. Однако от младенца этого ждать не приходится. Наконец, в процессе развития изменяется соотношение продолжительности каждого из состояний. Самое очевидное (хотя и не единственное), что происходит, — это резкое увеличение с возрастом времени бодрствования. Новорожденный спит в среднем 16–17 часов в сутки, а большую часть периода бодрствования у него занимает кормление. На пассивное бодрствование у новорожденного уходит только 10 % времени (Berg, Adkinson, & Strock, 1973; Berg & Berg, 1987).

О значении, которое имеет состояние, мы уже упоминали. Для понимания поведения ребенка необходимо знать, в каком он находится состоянии. В некоторых случаях состояние выступает в качестве одной из переменных, на выявление эффекта которых направлено исследование. Исследователей, к примеру, интересуют различия в реакции на тактильную стимуляцию в состоянии сна и бодрствования (например, Berg, Berg, & Graham, 1971). В некоторых исследованиях состояние выступает в качестве зависимой переменной, и здесь цель состоит в выявлении факторов, влияющих на состояние. К примеру, проведены исследования, направленные на изучение результатов экспериментального вмешательства во время сна на последующее состояние малыша (например, Anders & Roffwarg, 1973), или влияния различного рода стимуляции на стабильность состояния (например, Edelman, Kraemer, & Korner, 1982). Однако чаще всего в исследованиях младенцев цель состоит не в изучении переменной состояния, а в ее *контроле*. Иными словами, исследователь стремится работать с малышами, находящимися в одном состоянии, вне зависимости от того, какое состояние считается наиболее подходящим для изучения интересующей исследователя проблематики. В большинстве случаев исследователь предпочитает работать с малышами, находящимися в состоянии спокойного бодрствования, поскольку оно является наиболее благоприятным для перцептивной и познавательной деятельности.

Пытаясь застать малышей в состоянии спокойного бодрствования, исследователи прибегают к разнообразным хитростям. Очень часто родителей просят приносить детей на обследование в то время, когда они обычно не спят и не капризничают. Периодов дневного сна, естественно, нужно избегать; однако существуют определенные разногласия по вопросу о том, каков оптимальный интервал между временем обследования и кормлением (Field, 1982). Кроме того, в экспериментальных условиях возможны разнообразные манипуляции. На состояние младенца влияет его положение. К примеру, было обнаружено, что, находясь в вертикальном положении, малыши более внимательны, чем в положении лежа (Korner & Thoman, 1970). Давно известно, что изменением положения тела можно повысить активность дремлющего или успокоить капризничающего малыша. Для успокоения капризничающих малышей используются также соски, однако есть основания полагать, что использование соски может отразиться и на других аспектах поведения малыша, поэтому этим средством не следует злоупотреблять (Field, 1982).

Но как бы ни был изобретателен исследователь, он все еще зависит от прихоти младенца. Такая зависимость от капризов малыша имеет определенные методологические следствия. Периоды обследования должны быть достаточно короткими, иначе младенец выйдет из желательного состояния. Поэтому может возникнуть необходимость разделить период обследования на несколько сеансов или использовать не внутрисубъектный, а междисциплинарный план. При использовании внутри-

субъектного плана контрбалансировка заданий или экспериментальных условий приобретает еще большее значение, чем обычно. Предположение, что порядок следования заданий не важен, сомнительно при изучении любой возрастной группы; однако при изучении младенцев, с лабильностью их состояния, об этом предположении не может быть и речи.

Наиболее серьезное методологическое следствие касается избирательности исследования младенцев. Эта избирательность действует на двух уровнях. Как мы уже отметили, избирательность существует в отношении того, кто останется в исследовании, а кто нет. Целиком обследование проходят только те малыши, у которых поддерживалось желательное состояние. Кроме того, существует избирательность в отношении того, какие именно аспекты поведения этих оставшихся испытуемых подвергаются анализу. Как правило, исследователь работает с малышом, который не спит, спокоен и всем доволен. Однако в таком состоянии малыши, особенно самые маленькие из них, находятся не слишком долго. Обычно исследование направлено на оценку *оптимального* уровня реагирования младенца, что вполне оправдано и важно. Тем не менее необходимо помнить, что *обычный* уровень реагирования малыша может не достигать этого оптимума.

Оценка реагирования

Контроль состояния — одна из двух больших трудностей исследования младенцев. Вторая из них — выявление реакций, по которым можно было бы судить о возможностях малыша.

Рассмотрим самый трудный случай: исследование новорожденных. Новорожденный не совсем беспомощен; он появляется на свет, уже обладая разнообразными адаптивными рефлексам, то есть врожденными произвольными реакциями на определенного рода стимулы. Однако новорожденный еще не способен производить сложные целенаправленные действия. Не владеет он, конечно, и речью. Последнее сохраняется на протяжении всего периода младенчества; в действительности, слово *infant* («младенец») означает «не умеющий говорить». Стоит задуматься на мгновение о том, в какой мере наша оценка старших детей и взрослых зависит от использования языка. С младенцами все вербальные методики, которые мы обычно применяем к испытуемым старших возрастных групп, исключаются. Как же тогда узнать, о чем думает или что ощущает малыш?

В значительной мере история развития исследований периода младенчества — это история постепенного открытия все новых и новых реакций, по которым можно судить о том, как младенец воспринимает окружающий его мир. Большая часть этой главы посвящена рассмотрению именно этих реакций. В этом же разделе дается предварительное описание некоторых из наиболее информативных и широко используемых методик оценки реагирования младенцев. Хотя на первый взгляд может показаться, что рассматриваемые реакции различаются, в действительности они имеют ряд общих характеристик: все они требуют от младенца минимальных моторных усилий, и поэтому все их можно наблюдать уже у самых маленьких детей; все реакции можно точно измерить; и все реакции (по крайней мере, во многих случаях) поддаются однозначной интерпретации.

Одна из часто используемых методик — оценка *зрительных фиксаций*. Даже новорожденный может смотреть и смотрит на предметы. Более того, даже новорожденный может демонстрировать некоторую избирательность в том, на что он смот-

рит. Как мы вскоре убедимся, именно благодаря тому факту, что младенцы на одни объекты смотрят дольше, чем на другие, ученые получили возможность изучать перцептивные способности и перцептивные предпочтения. Зрительные реакции также сыграли определенную роль в исследовании двух других вопросов, которые рассматриваются в этой главе: когнитивного и социального развития в раннем возрасте.

Вторая реакция, которую часто используют исследователи, — это *сосание*. Сосание также является врожденной реакцией и наблюдается у всех младенцев. В действительности, у большинства младенцев на сосание пустышки уходит больше времени, чем на сосание в период кормления. Обычно при измерении реакции сосания оценивается изменение частоты сосания в ответ на изменения в окружающей ребенка среде. Например, прекращает ли младенец сосать при внезапном появлении в поле зрения нового объекта?

Физиологические реакции составляют третью из важнейших категорий зависимых переменных. Здесь, очевидно, от малыша не требуется никаких моторных усилий. С помощью современных технологий можно измерять широкий диапазон физиологических реакций. Как правило, при исследовании младенцев измеряется частота сердечных сокращений. Параметр частоты сердечных сокращений, с точки зрения исследования, обладает рядом желательных качеств. В отличие от некоторых физиологических реакций, изменение частоты сердечных сокращений можно зафиксировать в ответ на любого рода стимул и в любой модальности. Физиологическая система, опосредующая изменение частоты сердечных сокращений, довольно хорошо развита уже с рождения, что делает возможным использование параметра частоты сердечных сокращений при работе даже с новорожденными. Наконец, изменение сердечной активности происходит в двух направлениях, то есть частота сокращений либо повышается, либо снижается. Как мы увидим, есть данные, свидетельствующие о том, что снижение частоты сердечных сокращений указывает на нечто отличное от того, на что указывает ее повышение. Поэтому с помощью измерения частоты сердечных сокращений можно получить информацию, недоступную при измерении реакций, имеющих одно направление.

Три рассмотренные выше реакции ни в коей мере не исчерпывают поведенческие проявления младенцев, представляющие интерес для специалистов в области психологии развития. Это краткое введение не дает, конечно, представления о том, как в действительности в исследованиях используются реакции фиксации взгляда, сосания или изменения частоты сердечных сокращений. Более подробно методики оценки реагирования будут описаны при рассмотрении различных типов исследования периода младенчества.

Возрастные сравнения

Вопрос возрастных сравнений в общем виде уже рассматривался в главе 3. Обратимся теперь к проблемам возрастного сравнения при изучении младенцев. Когда испытуемые — это младенцы, произвести валидное возрастное сравнение в чем-то легче, а в чем-то труднее. Легкость или трудность зависит как от конкретных возрастов, так и от конкретных оцениваемых реакций.

Рассмотрим сначала сравнение разных возрастных групп в рамках периода младенчества. Такие исследования, как и большинство исследований в области психологии развития, проводятся преимущественно методом поперечных срезов, а не в лонгитуде. Однако при изучении младенцев некоторые из недостатков планов поперечных срезов, о которых говорилось в главе 3, имеют не столь большое значение. В частности, вероятность эффектов когорты — то есть эффектов, обусловленных различиями между поколениями, а не по возрасту, — намного меньше, чем при сравнении выборок старшего возраста. За редкими исключениями исследователь может допустить, что 4-, 8- и 12-месячные малыши принадлежат к одному поколению и что любые из обнаруженных возрастных различий не смешиваются с различиями между когортами. Конечно, все еще важно определить, что является причиной этих возрастных различий, однако большая категория объяснений уже исключается.

Результаты лонгитюдных исследований тоже, как правило, лучше поддаются интерпретации, если испытуемые — младенцы. Вспомним, что главная проблема лонгитюдных планов — трудность разграничения эффектов возраста и времени проведения обследования. Это смешение сохраняется в любой возрастной группе, однако в исследованиях младенцев оно редко оказывается существенным фактором. Вряд ли в 12 месяцев малыш будет вести себя иначе, чем в 6 месяцев, лишь потому, что одно обследование проводилось в марте 1995 года, а другое — в сентябре 1994 года. Непродолжительность временного интервала и особенности периода младенчества делают маловероятным влияние на поведение историко-культурных изменений. При изучении младенцев не так остро стоят и другие проблемы, типичные для лонгитюдных исследований. К примеру, вероятность того, что искажающим фактором окажется осознание собственного участия в исследовании при изучении младенцев меньше, чем при изучении старших детей или взрослых. Наконец, если отбросить трудности интерпретации, провести лонгитюдное исследование с младенцами просто легче с практической точки зрения. Как мы узнали, одно из препятствий при проведении лонгитюдных исследований — это то, что исследователь должен ждать, пока испытуемые достигнут интересующего его возраста: пока 10-летние испытуемые превратятся в 18-летних, 50-летние — в 70-летних и т. д. Исследователь периода младенчества имеет возможность наблюдать резкие изменения, происходящие буквально в течение нескольких месяцев.

До этого момента речь шла в основном об исследованиях, проводимых исключительно с младенцами. Рассмотрим теперь случай, когда исследователь не ограничивается периодом младенчества и пытается сравнить младенцев с детьми старшего возраста. Здесь, естественно, возникает больше проблем. С увеличением возрастного интервала повышается значение упомянутых выше факторов смешения. С расширением возрастного диапазона актуальнее также становятся практические трудности подбора равноценных испытуемых в исследованиях методом поперечных срезов и трудности сохранения испытуемых — в лонгитюдных исследованиях. Однако, наверное, самым значимым оказывается обсуждавшийся в главе 3 вопрос *эквивалентности измерения*. Для сравнения возрастных групп требуются психологические методики, подходящие для разных возрастов. Мы сможем проинтерпретировать сходство или различия в результатах, или оценить индивидуальную ста-

бильность или индивидуальное развитие только в том случае, когда наши методики будут измерять один и тот же конструкт у разных групп.

Масштабы проблемы эквивалентности измерения становятся ясны, если задуматься о том, что отличает младенцев от детей старшего возраста. Сравним, к примеру, 8-месячных и 8-летних детей. Типичное поведение и ситуации, в которых оно наблюдается, в этих группах совершенно разные. Если взять самое очевидное, значительная часть процесса решения задач и социального взаимодействия у 8-летних детей опосредована речью, чего нельзя сказать о младенцах. Улыбка 8-месячного ребенка обусловлена иными причинами и несет иной смысл, чем улыбка 8-летнего ребенка. Плач 8-месячного ребенка отличается от плача 8-летнего ребенка еще в большей мере.

Даже в рамках периода младенчества заметные изменения, связанные с развитием, могут создать трудности с эквивалентностью измерения. Например, при измерении частоты сердечных сокращений. Накоплен значительный объем данных, свидетельствующих о том, что новорожденные реагируют на появление нового объекта повышением, а младенцы — понижением частоты сердечных сокращений (Berg & Berg, 1987). У испытуемых из старшей возрастной группы учащение сердцебиения связано с чувством страха или оборонительной реакцией, а понижение частоты сердцебиения связано с состоянием внимания. Поэтому некоторые исследователи считают это изменение в период младенчества признаком перехода от реагирования на новизну по типу защиты к реагированию по типу внимания. Однако другие специалисты утверждают, что этот переход обусловлен не переключением с защиты на внимание, а созреванием физиологической системы, регулирующей частоту сердечных сокращений (Porges, 1979). Поэтому тот факт, что одну и ту же реакцию можно измерять в разных возрастных группах, не гарантирует того, что она будет означать то же самое.

Перейдем теперь от общих вопросов к рассмотрению конкретных видов исследований младенцев. Оставшаяся часть главы посвящена проблематике, на изучение которой чаще всего направлены исследования младенцев: восприятию, когнитивному и социальному развитию в младенчестве.

Восприятие в младенчестве

При изучении восприятия младенцев особый интерес представляют два общих вопроса. Первый — вопрос *перцептивных способностей*. Насколько развиты перцептивные способности малыша при рождении и как развиваются эти способности в первые два года жизни? Эта большая область исследования охватывает множество специфических вопросов. Насколько четко новорожденный видит предметы, что он различает, а что нет? Каков у новорожденного порог чувствительности к звуку, и какие звуки он может различить, а какие — нет? Обладает ли восприятие новорожденного или младенца константностью, и если нет, в каком возрасте она развивается? Любой вопрос, касающийся восприятия в целом, можно задать и с точки зрения психологии развития, начиная с вопросов о восприятии младенцев.

Второй вопрос — вопрос *перцептивных предпочтений*. Имея определенные перцептивные способности, на какие объекты направляет свое внимание младенец?

На что малышам интересно смотреть, что им интересно слушать, до чего интересно дотрагиваться? И почему они предпочитают именно это?

Прежде чем обсуждать конкретные методы, следует сделать несколько общих замечаний об исследовании восприятия. Во-первых, большинство исследований, направленных на изучение восприятия в младенчестве, проведены по плану поперечных срезов. Хотя особое внимание к этому методу, в принципе, могло бы привести к искажению результатов, по причинам, обсуждавшимся в предыдущем разделе, сомнительно, чтобы использование плана поперечных срезов действительно привело к ошибочным выводам о возрастных изменениях. Кроме того, большинство исследователей интересуют не индивидуальные различия, а общие закономерности развития. Такое качество, как константность восприятия, к примеру, в конечном счете появляется у всех детей, и чаще всего исследователя интересует, когда и как оно развивается, а не индивидуальные различия в темпах и характере развития. Как мы узнали из главы 3, одно из главных оснований для использования лонгитюдного метода — изучение стабильности индивидуальных различий во времени. Поскольку при разработке темы восприятия в младенчестве индивидуальные различия анализируются редко, использование лонгитюдного плана не дает никаких преимуществ¹.

Второй момент имеет отношение к континууму «лаборатория/полевые условия», о котором шла речь в главе 6. Подавляющее большинство исследований, направленных на изучение восприятия в младенчестве, проводится в лабораториях. Кроме того, создаваемые условия очень близки тому, что действительно понимается под словом лаборатория, — искусственно созданные, с высоким уровнем контроля, резко отличающиеся от «реальной жизни». Это, конечно, связано не с тем, что в естественных условиях восприятие отсутствует; восприятие — это процесс, пронизывающий всю психическую деятельность. Однако для изучения восприятия, как правило, требуется точная информация о стимуле и точное измерение едва заметных реакций на него; а этим требованиям отвечают обычно лишь лабораторные условия.

Последнее замечание касается двух, уже упоминавшихся затруднений при проведении исследований с младенцами: изменчивости состояния малыша и ограниченного репертуара реакций. Ни для какого другого аспекта исследований младенцев эти две проблемы не являются столь актуальными, как для исследования восприятия. Специалисты, занимающиеся восприятием, стремятся работать с малышами, находящимися в состоянии спокойного бодрствования, то есть в состоянии, оптимальном для перцептивной деятельности. Это стремление влечет за собой все практические трудности и избирательность, о которых говорилось выше. Проблема измерения реакций встает не только по причине ограниченного репер-

¹ Исключением из этого правила являются некоторые исследования, направленные на изучение развития восприятия глубины. Исследователей интересует вероятность того, что появление реакции страха на визуальный обрыв (который описывается в этой главе) связано с локомоторным опытом. Один из подходов к изучению этого вопроса — лонгитюдный: исследование одной и той же выборки в течение нескольких недель с целью определения того, более ли характерна реакция страха для детей, которые начинают ползать раньше или ползают больше других малышей. Как правило, обнаруживается, что это действительно так (Bertenthal, Campos, & Barret, 1984).

туара действий младенца, но и вследствие особенностей самого процесса восприятия: феномена внутриспсихического опыта, который далеко не всегда связан с наблюдаемым, измеримым поведением. Поэтому найти такие реакции, по которым можно было бы судить о перцептивном опыте младенца, может оказаться особенно трудной задачей.

Метод предпочтений

Метод предпочтений — это способ изучения зрительного восприятия в младенчестве, разработанный Робертом Фанцем (Robert Fantz, 1961). Вопрос, на который он призван отвечать, — вопрос зрительного различения: может ли младенец отличить один зрительный стимул от другого? Этот вопрос, очевидно, является ключевым при любой оценке зрительного восприятия малыша.

Одна из первых моделей аппарата Фанца изображена на рис. 10.1. Младенца помещают в кабинку для просмотра так, чтобы он мог смотреть на один из двух расположенных над ним объектов. Объекты устанавливаются друг от друга на таком расстоянии, что смотреть на оба объекта одновременно невозможно; для того чтобы направить свой взгляд на один или другой, необходимо слегка повернуть голову. Поворот головы — это движение, доступное даже новорожденному. Помимо поворота головы, единственное, что требуется от младенца, — это зафиксировать взгляд на одном из двух объектов. Зависимая переменная — время, которое младенец смотрит на один из стимулов. Как явствует из изображения на рис. 10.1, фиксацию может легко измерить наблюдатель, стоящий над кабинкой и отмечающий, какой из объектов отражается в глазах ребенка.

Обычно при использовании процедуры Фанца сравниваемые стимулы предъявляются несколько раз. Один стимул появляется равное количество раз с правой и с левой стороны. Вопрос состоит в том, будет ли младенец смотреть на один стимул дольше, чем на другой. Если младенец действительно смотрит дольше на один из объектов, говорят, что он демонстрирует предпочтение этого объекта. А если младенец проявляет предпочтение, делают вывод о том, что, скорее всего, он различает предъявленные объекты. Логика очень проста: предпочтение могло возникнуть только при условии, что ребенок увидел разницу. Если стимулы воспринимаются как одинаковые, нет основания смотреть на один из них дольше, чем на другой.

Метод предпочтений применялся с множеством разнообразных зрительных стимулов. Например, его использовали для определения того, видит ли младенец разницу между фигурой, заполненной орнаментом, и равноценной однотонной фигурой (к примеру, между глазом и простым кругом), между цветным и равноценным черно-белым объектом, между плоским и объемным изображением одного и того же объекта, между новым и знакомым объектом, между лицом матери и лицом постороннего человека. Заметьте, что этот метод позволяет не только получить информацию о способности к различению, но и ответить на поставленный выше вопрос о перцептивных предпочтениях. Предположим, мы обнаружили (и это соответствует действительности), что новорожденный на заполненные орнаментом фигуры смотрит дольше, чем на однотонные. Этот результат свидетельствует о способности малыша отличать заполненную орнаментом фигуру от одно-

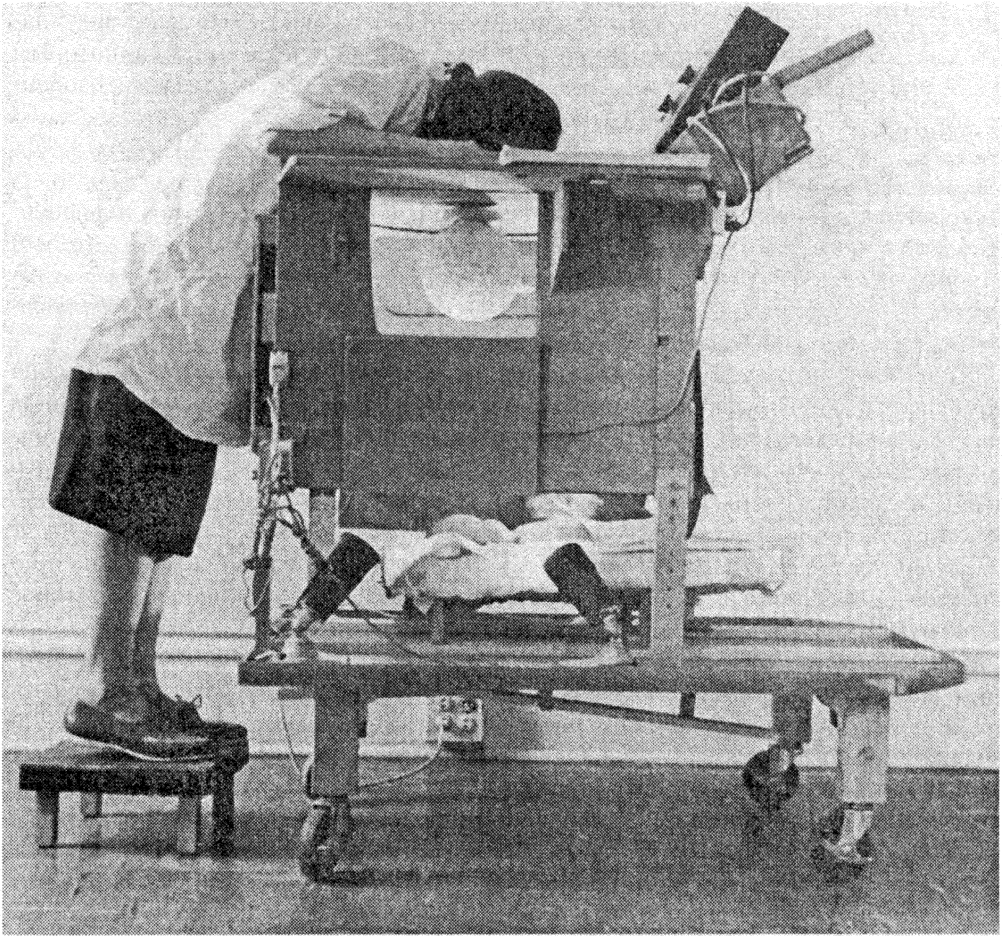


Рис. 10.1. Аппарат для использования метода предпочтений Фанца при изучении восприятия младенцев. (Источник: «The Origin of Form Perception», R. L. Fantz, 1961, *Scientific American*, 204, p. 66. Фото: David Linton)

тонной. Кроме того, он свидетельствует о том, что ребенок не просто видит рисунок внутри контура, но и *предпочитает* на него смотреть.

Преимущества метода предпочтений заключаются в минимальных требованиях к реакциям младенца и возможности широкого применения. Однако этот метод имеет и некоторые недостатки. Иногда бывает трудно установить, что именно различает ребенок. Если малыш демонстрирует предпочтение, мы знаем, что он видит разницу между объектами, но мы не можем знать наверняка, по каким признакам объекты воспринимаются как различные. Один из способов разрешения этой проблемы — попытаться подобрать такие объекты, которые отличались бы только по одному признаку, например круг и шар тех же размеров и того же цвета. Если

том стойкого позиционного предпочтения будет равное время фиксации взгляда на обоих объектах, а следовательно, и видимое отсутствие предпочтений. Этот случай прост в том смысле, что систематическое искажение реакции можно легко обнаружить и тем самым исключить возможность ошибочного вывода о способности или неспособности младенца различить объекты. Однако такое искажение оставляет неразрешенным вопрос, способен ли ребенок к различению.

В более сложном случае младенец не демонстрирует систематического искажения в реакциях, а равномерно распределяет свое внимание между двумя объектами. Это может быть обусловлено неспособностью младенца увидеть разницу между объектами; как отмечалось ранее, предпочтение возможно только при различении. Однако этот негативный вывод нельзя сделать со всей определенностью; возможно, младенец отлично видит разницу между объектами, но просто равнодушен к ним обоим. Как выразил эту мысль Флэйвелл (Flavell, 1985), «из предпочтения логически следует различение, однако из различения определенно не следует предпочтение» (р. 170). И это — основное, что вызывает затруднения при интерпретации результатов использования метода предпочтений: негативные результаты могут означать как отсутствие различения, так и просто отсутствие предпочтений.

Габитуация/дегабитуация

Представьте себе следующую последовательность событий. Вы входите в незнакомую комнату, и сначала вас отвлекает громкое тиканье часов. Вы садитесь и принимаетесь за чтение книги, проходит несколько минут, и вы уже совсем не замечаете тиканья. Затем часы перестают тикать, и вы тут же реагируете на это изменение.

Эта последовательность событий иллюстрирует три взаимосвязанных психологических процесса, имеющих огромное значение. При появлении нового стимула многие существа, включая и человека, демонстрируют то, что называется **ориентировочной реакцией**. Ориентировочная реакция — это на самом деле комплекс связанных между собой реакций, выполняющих функцию максимизации внимания к новому событию. Обычно эти реакции включают в себя мгновенное прерывание текущей деятельности, ориентирование рецепторов в направлении нового стимула и ряд характерных физиологических изменений (например, снижение частоты сердечных сокращений). Этот набор реакций прост, но его адаптивное значение весьма велико, поскольку он представляет собой средство автоматического направления внимания на все новое, что появляется в среде.

Рассмотрим теперь, что же происходит при многократном действии стимула, вызвавшего изначально ориентировочную реакцию, как в примере с тикающими часами. В конечном счете ориентировочная реакция ослабнет или совсем угаснет. Это угасание ориентировочной реакции при продолжительном действии стимула называется **габитуацией**. Габитуация — также важнейший адаптивный процесс: как только стимул перестает быть новым, пристальное внимание к нему теряет свою необходимость.

Рассмотрим, наконец, что происходит при изменении стимула, к которому возникло привыкание, например при прекращении тиканья. Это изменение в привычной стимуляции вызовет новую реакцию внимания. Возобновление внимания при

изменении стимула, вызвавшего привыкание, называется **дегабитуацией**. Дегабитуация также выполняет адаптивную функцию: при изменении стимуляции вновь возникает необходимость ориентировки на новизну.

Как можно предположить на основании того, что описание трех этих феноменов было включено в данную главу, все они наблюдаются в младенческом возрасте. С какого возраста их можно обнаружить — вопрос спорный; но сегодня уже ясно, что даже у новорожденных при оптимальных для этого условиях можно вызвать ориентировочную реакцию, габитуацию и дегабитуацию (Berg & Berg, 1987). Обычно в качестве зависимой переменной в исследованиях младенцев выступают изменения частоты сердечных сокращений: снижение ее считается одним из признаков ориентировочной реакции, а слабое изменение или полное его отсутствие — одним из признаков габитуации. Оцениваются также и две другие описанные выше реакции. При измерении интенсивности сосания малыш сосет соску со специальным датчиком; прекращение сосания рассматривается как свидетельство ориентировочной реакции, а непрерывное сосание — как свидетельство габитуации. При измерении длительности зрительных фиксаций главным является общее время фиксации взгляда на стимуле; изначально пристальное внимание указывает на ориентировочную реакцию, а снижении времени фиксации — на габитуацию.

Как можно использовать эти процессы для изучения восприятия в младенчестве? На элементарном уровне ориентировочную реакцию можно использовать для ответа на вопрос о распознании: может ли младенец вообще распознать стимул? Очевидно, что если уровень интенсивности стимула находится ниже порога чувствительности для определенной модальности восприятия, ориентировочная реакция на этот стимул не возникнет. Это кажется чем-то само собой разумеющимся, однако поведенческие реакции, которые можно рассматривать как признак распознавания, зачастую едва заметны. В частности, иногда физиологические реакции свидетельствуют о том, что младенец распознал стимул, в то время как никакой внешней реакции на этот стимул не наблюдалось вообще.

Допустим, нас интересует не просто распознавание единичного стимула, а восприятие разницы между двумя стимулами. Здесь также вполне уместно использовать последовательность «габитуация—дегабитуация». Предположим, мы хотим узнать, может ли младенец различить два звуковых сигнала, скажем, слог «па» от слога «ба». Для этого мы непрерывно даем сигнал «па», пока младенец к нему не привыкает. В тот момент, когда должно последовать очередное «па», мы меняем его на «ба». Если у младенца наблюдается реакция дегабитуации, это говорит о том, что он различает эти слоги.

Этот пример демонстрирует очевидные преимущества метода габитуации перед методом предпочтений: более широкие возможности применения. Метод предпочтений нельзя применить для изучения слухового восприятия; довольно трудно, дав одновременно два звуковых сигнала, узнать, какому из них ребенок отдал предпочтение. Однако довольно легко можно давать один сигнал, пока не исчезнет реакция, а затем дать другой звуковой сигнал, отличающийся по некоторому признаку от первого. В действительности габитуация — самый популярный метод

изучения слухового восприятия младенцев. Этот метод применяется также и при изучении обонятельного восприятия, то есть способности младенца различать запахи (например, Engen, Lipsitt, & Kaye, 1963). Кроме того, он используется для изучения зрительного различения.

Обусловливание

Габитуацию можно рассматривать как своего рода научение — научение не реагировать на знакомый стимул. В этом разделе мы поговорим еще об одной из основных форм научения — о научении повторять действие, подкрепляемое наградой. Научение, основанное на получении подкрепления, называется *оперантным* (или *инструментальным*) *обусловливанием*.

При изучении восприятия в младенчестве оперантное обусловливание используется несколькими способами. Простейший из них имеет много общего с описанной выше парадигмой габитуации/дегабитуации. Рассмотрим применение оперантного обусловливания при изучении восприятия речи. Начнем исследование с того, что будем подкреплять сосание ребенком соски; при этом интенсивность подкрепления будет пропорциональна интенсивности сосания. Чем быстрее будет сосать ребенок, тем чаще он будет получать подкрепление. Такая ситуация называется «сопряженным подкреплением». Подкрепление будет заключаться в том, что будет произноситься какой-либо звук речи, скажем слог «ба». Сначала это подкрепление довольно эффективно, и младенец сосет быстро. Однако затем «ба» начинает терять свою привлекательность и частота сосательных движений снижается. Такое снижение является отражением *пресыщения* — падения ценности подкрепления как функции от многократного подкрепления. По достижении реакции сосания низкого уровня мы меняем подкрепление «ба» на «па». Если интенсивность сосания существенно возрастет, мы сможем сделать вывод о том, что младенец отличает «па» от «ба».

При втором способе обусловливания, который называется **обусловленным поворотом головы**, используется способность ребенка научиться связывать определенный стимул с некоторой реакцией. Предположим, что справа от малыша подается звуковой сигнал. Сначала малыш, скорее всего, будет оборачиваться на звук; однако если этот сигнал подавать многократно, ориентировочная реакция угаснет. Но что произойдет, если добавить подкрепление? Предположим, что всякий раз, когда ребенок будет поворачиваться в ответ на звуковой сигнал, будет появляться интересный зрительный стимул, например мерцающие огоньки или движущаяся игрушка. В этом случае реакция поворота головы не будет ослабевать, а звуковой сигнал станет командой, или *различительным стимулом*, к реакции: каждый раз, когда малыш будет слышать звуковой сигнал, надеясь получить подкрепление, он будет поворачивать голову направо. Как только эта связь закрепится, стимул можно будет всячески изменять с тем, чтобы оценить возможности слухового восприятия малыша. Для определения порога слуховой чувствительности ребенка мы могли бы изменять интенсивность или частоту звукового сигнала (например, Schneider, Trehub, & Bull, 1980). Или же мы могли бы предъявлять два разных

стимула, только один из которых был бы связан с подкреплением. Поворот головы в ответ лишь на один из стимулов свидетельствовал бы о том, что малыш воспринимает разницу между стимулами.

Третий подход, основанный на обусловливании, сходен с описанной только что процедурой, однако здесь, кроме всего вышесказанного, делается акцент на генерализации. *Генерализация* — это перенос реакции, вызванной одним стимулом, на другие, похожие на предыдущий стимулы. Поскольку степень генерализации зависит от уровня сходства между исходным и новым стимулом, генерализацию можно использовать для оценки того, насколько сходными воспринимает малыш некие два стимула.

В качестве примера рассмотрим эксперимент Бауэра (Bower, 1966). Исследование Бауэра было направлено на изучение одного из классических вопросов теории восприятия: константности. Константность действительно является вопросом для теории восприятия, поскольку мы воспринимаем окружающее лучше, чем должны были бы воспринимать, учитывая то, какое изображение мы получаем на сетчатке. К примеру, нам как-то удается воспринимать вещи обладающими постоянными размерами, несмотря на непрерывное изменение величины образа на сетчатке; обладающими постоянной формой — несмотря на непрерывное изменение ориентации предмета в пространстве и т. д. Конкретный вопрос, к изучению которого обратился Бауэр, касается константности величины объектов в восприятии младенцев. Видит ли младенец мир, так же как и мы, состоящим из предметов, которые обладают постоянными размерами, или предметы в восприятии младенца увеличиваются (уменьшаются) с увеличением (уменьшением) образа на сетчатке?

Бауэр начал с того, что вызвал у своих испытуемых (младенцев в возрасте 60–80 дней) простую обусловленную реакцию поворота головы для получения подкрепления. Как и в только что рассмотренных исследованиях, обусловливание было проведено с использованием различительного стимула. В исследовании Бауэра различительным стимулом был куб со стороной 30 см, появлявшийся перед младенцем на расстоянии 1 м. Повороты головы подкреплялись только при появлении куба; при его отсутствии они не получали подкрепления. То есть на первом этапе исследования младенцы научались поворачивать голову всякий раз, когда куб появлялся перед ними на строго определенном расстоянии.

На втором этапе исследования проводилась проверка генерализации. Каждому младенцу в определенном порядке предъявлялось по 4 стимула. Характеристики этих стимулов представлены в табл. 10.2. Если бы у младенцев восприятие величины зависело бы от образа на сетчатке, они должны были бы реагировать на стимул 4 так же, как и на стимул 1. В действительности, как явствует из таблицы, к стимулу 4 отношение было совершенно иное, чем к стимулу 1, что выражалось наиболее низким уровнем реагирования. Однако стимул 2 — обладавший такими же размерами, как и стимул 1 — вызвал довольно высокий уровень реагирования, хотя и находился на ином расстоянии и поэтому по-другому отражался на сетчатке. Именно на этом основании Бауэр и заключил, что уже 2-месячные младенцы обладают способностью константно воспринимать величину объекта.

Должен отметить, что эксперимент Бауэра воспроизводился в дальнейшем с переменным успехом. Однако для нас скорее важна методология, а не то, верными

или ошибочными оказались выводы о константности. Как и габитуация, генерализация представляет собой простой, но довольно эффективный метод определения того, что для младенца выглядит одинаковым, а что — нет. Как и подход, основанный на габитуации, подход, основанный на генерализации, можно использовать для изучения многих аспектов восприятия.

Таблица 10.2

Стимулы и результаты фазы генерализации в проведенном Бауэром исследовании константности восприятия величины объектов

Стимул		Связь с различительным стимулом	Среднее количество реакций
Размеры	Расстояние		
30 см	1 м	Идентичен	98
30 см	3 м	Тех же размеров	58
90 см	1 м	На том же расстоянии	54
90 см	3 м	Тот же образ на сетчатке	22

Источник: «The Visual World of Infants», G. R. Bower, 1966, Scientific American, 215, 90–92.

Визуальный обрыв

Последняя процедура, которую мы рассмотрим, имеет не такие широкие возможности для применения, как описанные выше методы. Метод **«Визуального обрыва»** (разработанный Уолком и Гибсоном (Walk & Gibson, 1961)) используется для получения ответа на один вопрос: воспринимают ли малыши глубину? С помощью этого метода оценивается вполне конкретная форма восприятия глубины — восприятие перепада высоты, или «обрыва». Например, воспринимает ли младенец перепад высоты, высовываясь из кроватки и смотря на пол? Этот вопрос имеет как теоретическое, так и практическое значение.

«Визуальный обрыв» изображен на рис. 10.3. Как можно заметить, это приспособление состоит из большого стола со стеклянной крышкой, разделенной на две половины. При использовании этого метода под стекло обычно кладут клетчатую ткань. Одной стороной она непосредственно примыкает к стеклу, создавая впечатление твердой поверхности. Эту часть называют «мелкой». На другой стороне ткань располагается на несколько футов ниже уровня стекла, создавая впечатление перепада высоты или обрыва. Эту часть называют «глубокой». Как правило, проверка начинается с того, что младенца в положении на четвереньках ставят на проходящую по центру стола границу. Затем мать подходит с разных сторон и пытается уговорить младенца подползти к ней по стеклу. Интерес представляет дифференцированная реакция: поползет ли младенец по «мелкой» стороне и откажется ли он ползти по «глубокой» стороне? Если произойдет именно это, можно будет сделать вывод о том, что младенец воспринимает глубину.

Исследования такого рода показали, что уже в 6-месячном возрасте младенцы часто (хотя и не всегда) стараются избегать «глубокой» стороны, тем самым демонстрируя определенную способность воспринимать глубину (Bertenthal & Campos,

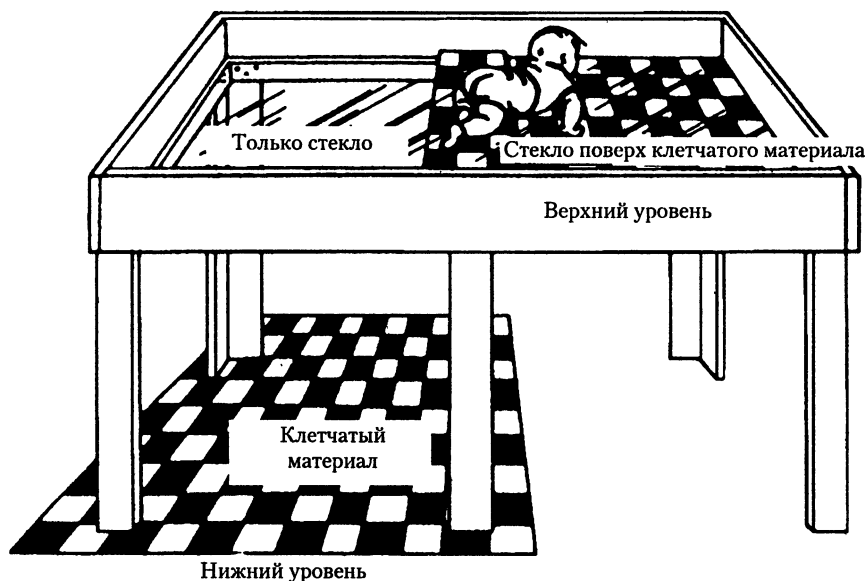


Рис. 10.3. Приспособление «Визуальный обрыв» для изучения восприятия глубины.
(Источник: «A Comparative and Analytic Study of Visual Depth Perception», R. D. Walk & E. J. Gibson, 1961, *Psychological Monographs*, 75.)

1990). Детей младше 6-месячного возраста таким способом проверять нельзя, поскольку результаты этого метода зависят от умения ребенка ползать. Таким образом, классическая процедура использования «Визуального обрыва» не позволяет определить, с какого возраста возможно восприятие глубины. Альтернативный способ проверки, разработанный специально для самых маленьких, состоит в том, что младенцу просто дают посмотреть сверху вниз с обеих сторон «обрыва» и оценивают его реакцию. Обычно измеряются такие переменные, как плач или изменение частоты сердечных сокращений. Используя этот подход, исследователи обнаружили, что уже в 2-месячном возрасте дети неодинаково реагируют на разные стороны «обрыва», таким образом, демонстрируют способность видеть разницу между этими сторонами. Однако у них не наблюдается ярко выраженных реакций страха на «глубокую сторону» и поэтому их способность воспринимать глубину остается под вопросом (Campos, Langer, & Krowitz, 1970).

«Визуальный обрыв» — не единственная процедура проверки восприятия младенцами глубины. Среди других методик можно назвать оценку реакции протягивания малышом руки, чтобы достать предметы, находящиеся на разном расстоянии (например, Field, 1977); и оценка реакции избегания на стремительно приближающиеся объекты (например, Younas, 1981). В целом, в своем обзоре методов изучения восприятия в младенчестве мы лишь слегка затронули эту широкую и быстро развивающуюся область исследования. Более подробное рассмотрение этой темы можно найти в работах Aslin, Pisoni, & Jusczyk, 1983; Gottlieb & Krasnegor, 1984; Fox & Fitzgerald, 1990; Haith, 1990.

Когнитивное развитие в младенчестве

Хотя в последнее время подход Пиаже утратил свое господствующее положение, он остается наиболее значимым из подходов к изучению когнитивного развития в младенчестве. Поэтому в этом разделе основное внимание будет уделено работам Пиаже (Piaget, 1951, 1952, 1954) и некоторым из его последователей. В следующей главе, в которой речь идет о когнитивном развитии в целом, я коснусь некоторых других подходов к изучению когнитивной сферы младенцев.

Исследования периода младенчества, проведенные Пиаже, слишком обширны, чтобы уместить их описание в одном разделе. Здесь я ставлю перед собой лишь две цели. Во-первых, изложить некоторые общие аспекты пиажетианского подхода к изучению младенцев. И, во-вторых, описать процедуры, используемые для изучения одного из наиболее важных достижений младенчества — понятия об объекте.

В основе выводов Пиаже о развитии в период младенчества лежат его собственные исследования своих троих детей в первые 2–3 года их жизни. Его подход — это разработка одного из исторически самых ранних методов детской психологии, метода детской биографии. Используя **метод детской биографии**, родитель-ученый внимательно наблюдает за развитием собственного ребенка, записывает свои наблюдения, которые затем составляют базу данных для получения некоторых обобщенных выводов о развитии человека. Такие исследования стали довольно популярны во второй половине XIX века. Наиболее известным предшественником Пиаже в этом деле стал Чарлз Дарвин (Darwin, 1877), изучавший своего маленького сына.

Если говорить о Пиаже, он не ограничился простым описанием заинтересовавшего его поведения. С самого начала в своей работе Пиаже руководствовался теорией, последовательно пытаясь ответить на важнейшие философские вопросы, например, об истоках представлений о пространстве, времени и причинно-следственных связях. Ниже мы рассмотрим одно из таких понятий, понятие о постоянстве объекта. Кроме того, работа Пиаже состояла не просто из наблюдения за естественным поведением. Наблюдение в естественных условиях *использовалось* — многие сотни часов довольно утомительного наблюдения. Однако эти наблюдения постоянно дополнялись попытками экспериментирования. Если, к примеру, однажды Пиаже интересовывала реакция его дочери на препятствия, он не ждал, пока что-нибудь преградит ей путь. Вместо этого он мог поставить между дочерью и ее любимой игрушкой подушку, а затем записать, как она справилась с этой трудностью.

Наблюдения Пиаже собраны в сотнях «протоколов», и, не процитировав некоторые из них, невозможно передать специфику этой работы. Вставка 10.1 содержит подборку протоколов, касающихся постепенного освоения младенцем принципов причинности. Заметьте, как естественное поведение переплетается с экспериментально вызванным. Заметьте также, как описание непосредственной данности сочетается с интерпретирующим комментарием.

Для того чтобы вычленить общие характеристики подхода Пиаже, полезно сравнить его исследования с работами по восприятию младенцев, о которых шла речь в предыдущем разделе. Можно заметить ряд отличий. Свои исследования Пиаже

Вставка 10.1

Примеры протоколов Пиаже, описывающих понимание младенцами причинно-следственных связей**Набл. 128**

В возрасте 0;3(12), то есть через несколько дней после того, как он научился хватать видимые предметы, Лоран оказывается перед подвешенной над его кроваткой погремушкой; с погремушки свисает цепочка от часов... Если говорить о связи между цепочкой и погремушкой, результат эксперимента однозначно отрицательный: Лоран не тянет за цепочку, и когда я кладу ее ему в руку, он, случайно дернув за нее и услышав звук, машет рукой, но выпускает цепочку. С другой стороны, он, по-видимому, тут же устанавливает связь между движениями своей руки и движениями погремушки, так как, случайно помахав рукой и услышав звук погремушки, он, глядя на погремушку, начинает махать пустой рукой вновь и вновь и даже машет все сильнее и сильнее...

Набл. 134

В возрасте 0;7(7) Лоран внимательно смотрит, как я барабаню пальцами по жестяной коробке размерами 15 × 20 см. Коробка лежит перед ним на подушке, на два сантиметра дальше радиуса досягаемости. Окончив свою игру, под взглядом Лорана я кладу свою руку на пять сантиметров от его руки. Пока я барабаню пальцами, Лоран блаженно улыбается; но когда я останавливаюсь, он какое-то время смотрит на мою руку, затем сразу начинает изучать коробку, а затем, глядя на коробку, хлопает в ладоши, машет обеими руками, качает головой, выгибается и т. п. Иными словами, использует весь свой арсенал магически-феноменистических приемов. Что касается моей руки, лежащей перед ним, он дважды хватает ее, трясет, машет ею и т. д. Однако он не подводит ее обратно к коробке, хотя это было бы легко сделать. Не пытается он и придумать какой-нибудь новый прием, чтобы вновь извлечь из нее звуки.

проводил в естественной для ребенка среде; исследования же восприятия, как мы увидели, практически полностью ограничены лабораторной обстановкой. В исследованиях Пиаже по существу не использовалось никакое специальное оборудование; ребенок взаимодействовал со знакомыми объектами и людьми, а единственным «регистрирующим устройством» был сам отец. Изучение восприятия, напротив, предполагает интенсивное использование сложной аппаратуры, например, камеры для слежения за движениями глаз или электрокардиографов. Выводы Пиаже зависели от наблюдений человека; для исследования же восприятия нормой является автоматическая регистрация. Наконец, исследования Пиаже были лонгитюдными, тогда как исследования восприятия в младенчестве построены преимущественно по плану поперечных срезов.

Почему так много различий в подходе? Эти различия частично — функция от изучаемой проблематики, а частично — следствие определенных особенностей исследовательского стиля Пиаже. Как было сказано в предыдущем разделе, феномены, изучаемые в исследованиях восприятия в младенчестве, невозможно измерить методом наблюдения в естественных условиях. Какое естественное и наблюдаемое поведение младенца позволит судить о том, воспринимает ли он удаляющийся предмет константным в своих размерах или изменяющимся? Поведенческие же проявления когнитивного развития, напротив, часто находят выражение в естественном взаимодействии ребенка со средой, особенно в младенчестве. Один из примеров — исследование младенцем причинно-следственных связей, зафиксиро-

Набл. 149

В возрасте 1;0(29) Жаклин впервые видит известную игрушку, состоящую из цыплят, приводимых в движение грузом. Цыплята размещены на деревянном кольце, и спереди каждый из них привязан к тяжелому шарiku, который находится ниже плоскости кольца; поэтому малейшее перемещение шарика приводит цыплят в движение, и они начинают ударять клювами по краю кольца.

Сначала Жаклин, немного рассмотрев игрушку, которую я привожу в движение легким покачиванием, дотрагивается до шарика и замечает одновременное движение цыплят. Затем она целенаправленно приводит шарик в движение, следя за цыплятами. Убедившись в наличии некой связи, природа которой ей, очевидно, не совсем понятна, Жаклин своим указательным пальцем начинает подталкивать шарик всякий раз, как игрушка останавливается.

В этом примере Жаклин не приписывает спонтанное движение шарiku (как она это делала в предыдущем примере с мячом или плюшевой игрушкой), а, несомненно, воспринимает движение шарика как причину движения цыплят. Таким образом, с этой точки зрения, имеет место объективизация причинно-следственной связи. Более того, для Жаклин шарик — это не элементарное продолжение ее руки (как тесемки, подвешенные над кроватью и т. д.); она приводит его в движение, просто отпуская.

Набл. 157

В возрасте 1;6(8) Жаклин сидит на кровати рядом с матерью. Я сижу в ногах кровати, на стороне, противоположной той, где сидит Жаклин, и она меня не видит и не знает, что я в комнате. Я поднимаю и начинаю покачивать над кроватью прутком с привязанной на конце кисточкой. Жаклин наблюдает с большим интересом. «Прутик, пруттик», — восклицает она и внимательно следит за раскачиванием. В определенный момент она прекращает смотреть на кончик прута и, по-видимому, пытается разобраться в происходящем. Затем она пытается взглянуть на другой конец прута и для этого, опершись на мать, оборачивается назад и замечает меня. Она не выражает никакого удивления, будто уже до этого знала, что причиной всему был я...

Источник: *The Construction of Reality in the Child* (pp. 261, 276, 309–310, 334), J. Piaget, 1954, New York: Basic Books, Inc.

ванное в протоколах из вставки 10.1. Еще один пример — изучение понятия об объекте, о котором мы вскоре поговорим более подробно. Хотя эта граница довольно зыбка (как и различие между «восприятием» и «познанием» в целом), есть основания утверждать, что познавательную деятельность, особенно в младенчестве, наблюдать значительно легче, чем перцептивную.

Второе основание для различий связано с излюбленным подходом Пиаже к исследованию. Как будет сказано в следующей главе, многие особенности исследований Пиаже, проводившихся с младенцами, характерны и для его работ со старшими детьми. К этим особенностям относится и предпочтение гибких приемов строго стандартизированным, предпочтение простых и знакомых материалов — сложной аппаратуре и предпочтение интерпретационного анализа отдельных протоколов — стандартным процедурам статистической обработки. Кроме того, некоторые особенности исследований Пиаже непосредственно вытекают из условий, в которых он работал. С выборкой, состоящей из трех человек, вряд ли можно произвести хоть что-то похожее на статистический анализ. Производство регистрации в домашних условиях налагает строгие ограничения на возможности использования специальной аппаратуры; и в любом случае, многие из технических приспособлений, которые являются неотъемлемой частью современного исследования, 65 лет назад, когда Пиаже начинал свою работу, еще не были изобретены.

Подход Пиаже к изучению младенцев имеет как сильные, так и слабые стороны. Недостатки даже, наверное, более очевидны. Выборка Пиаже мала и не отвеча-

ет требованию случайного отбора. Полностью полагаясь на данные, полученные методом наблюдения, Пиаже тем не менее не делал никаких попыток доказать надежность, основанную на общности оценок нескольких наблюдателей. Даже с выборкой в три человека есть возможность произвести стандартизацию и контроль, возможность, которая была упущена. Метод представления результатов стирает различие между данными и выводами. А чрезмерное внимание Пиаже к наблюдаемой моторной деятельности при диагностике развития в младенчестве может привести к недооценке возможностей ребенка. Вскоре мы еще вернемся к этим критическим замечаниям.

О достоинствах исследований Пиаже уже говорилось ранее. В главе 5, при рассмотрении эксплоративных исследований, работы Пиаже приводились в качестве яркого примера, иллюстрирующего ценность гибкого, ориентированного на открытие подхода к изучению ребенка. В главе 3, при рассмотрении лонгитюдных исследований, проведенные Пиаже исследования младенчества приводились в качестве примера, иллюстрирующего ценность всестороннего анализа отдельных случаев при изучении развития детей. В данной области, пожалуй, нет лучшего примера, демонстрирующего достоинства наблюдения в естественных условиях, проводимого искушенным наблюдателем. В целом, тот факт, что работы Пиаже, уже более 50 лет спустя после выхода их в свет, продолжают иметь влияние на исследования развития младенчества, — достаточное подтверждение их жизнеспособности.

Обратимся теперь к исследованию **понятия об объекте**. Термин *понятие об объекте* означает осознание того, что объекты существуют независимо от нашего перцептивного контакта с ними. То есть это осознание того, что объект не исчезает лишь из-за того, что в данный момент мы его не видим, не слышим, не чувствуем на ощупь и т. п. Трудно представить себе нечто более элементарное из того, что необходимо осознавать. Однако на основе своих исследований Пиаже сделал вывод о том, что на момент рождения понятие об объекте отсутствует и лишь постепенно развивается в течение первых двух лет жизни; а также, что, овладевая этим понятием, младенец проходит ряд определенных стадий. Эти положения концепции Пиаже имеют огромное теоретическое и практическое значение, что объясняет то внимание, которое было уделено их проверке в дальнейших исследованиях. Для нас же наибольший интерес представляет методологический аспект проблемы. Как Пиаже — или любой другой исследователь — определяет, что младенец знает об объектах, а что нет?

Как отмечалось, наблюдения Пиаже зафиксированы в форме протоколов — если говорить о понятии об объекте, то в 66 протоколах, большая часть которых содержит множество наблюдений. Некоторые из них представлены во вставке 10.2. Эти выборочные протоколы иллюстрируют замечания, которые будут сделаны ниже.

В своих исследованиях понятия об объекте Пиаже ориентировался в основном на реакцию ребенка при исчезновении предметов. Такие исчезновения, разумеется, довольно часто происходят в ходе естественного взаимодействия ребенка с объектами. Игрушка скатывается с высокого стульчика или выпадает из кровати. Мать или отец выходит из комнаты. Малыш осматривается, но уже не видит рожок или маму. Помимо того, что эти события естественны, экспериментатор довольно легко сам может создать определенную ситуацию исчезновения. Папа Пиа-

Вставка 10.2

Примеры протоколов Пиаже, описывающих понятие младенца об объекте**Набл. 2**

Уже в возрасте 0;2(27) Жаклин следит взглядом за своей матерью и, когда та выходит за пределы зрительного поля, продолжает смотреть в том же направлении, пока мать не появится вновь.

То же наблюдение с Лораном в возрасте 0;2(1). Я смотрю сверху на лежащего в кроватке Лорана и время от времени появляюсь примерно в одном и том же месте; когда меня не видно, Лоран смотрит в эту точку, очевидно ожидая, что я появлюсь вновь.

Набл. 28

В возрасте 0;7(28) Жаклин пытается достать целлулоидного утенка, поставить его на смятое одеяло. Но утенок скатывается и оказывается очень близко к руке Жаклин, но за складкой простыни. Жаклин следит за утенком, даже повторяя рукой траекторию падения игрушки. Но как только утенок исчезает — все! Жаклин не приходит в голову поискать утенка в складках простыни, что было бы сделать совсем нетрудно...

Затем я достаю потерявшегося утенка и трижды кладу его около руки Жаклин. Все три раза она пытается его схватить, но всякий раз, когда она уже почти касается ее, я вновь на ее глазах прячу утенка под простыней. Жаклин тут же убирает руку и теряет интерес. Во второй и третий раз я зажимаю утенка, укрытого простыней, в ее руке; Жаклин некоторое время машет утенком, но ей не приходит в голову откинуть простыню.

Набл. 40

В возрасте 0;10(18) Жаклин сидит на матрасе, ее ничто не тревожит и не отвлекает (никаких покрывал и т. п.). Я беру из ее рук игрушечного попугая и дважды прячу его под матрас, слева от нее, в место А. Оба раза Жаклин немедленно начинает поиски игрушки и находит ее. Затем я беру из ее рук попугая и медленно на ее глазах переносу его под матрас справа от нее, в место В. Жаклин внимательно следит за этим перемещением, но в тот момент, когда попугай исчезает под матрасом в месте В, она поворачивается налево и пытается найти его там, где он был спрятан до этого, в А.

Следующие четыре раза я прячу попугая только в В, не пряча его предварительно в А. Каждый раз Жаклин очень внимательно следит за мной. Тем не менее каждый раз она пытается найти игрушку в А; она подворачивает матрас и целенаправленно обыскивает его. Однако во время последних двух попыток радиус поиска уменьшается.

Набл. 55

В возрасте 1;8(8) Жаклин сидит на зеленом коврике и играет с картофелиной, которая вызывает у нее живейший интерес (это новый для нее предмет). Она произносит «ка-тооша» и развлекается тем, что кладет ее в пустую коробку, а затем вынимает ее оттуда. Несколько дней она в полном восторге от своей игры.

Я беру картофелину и на глазах у Жаклин кладу ее в коробку. Затем я прячу коробку под коврик, незаметно для ребенка переворачиваю ее там и вытаскиваю пустой. Жаклин все это время непрерывно смотрела на коврик и осознала, что я под ним что-то сделал. Я прошу: «Дай папе картошку». Она пытается найти картошку в коробке, смотрит на меня, затем вновь внимательно осматривает коробку, бросает взгляд на коврик и т. д., но ей не приходит в голову откинуть ковер, под которым лежит картофелина.

Все следующие пять попыток не приводят ни к какому результату. Однако я пробую вновь. Каждый раз на глазах у ребенка я помещаю предмет в коробку, кладу коробку под коврик и достаю ее оттуда пустой. Каждый раз Жаклин заглядывает в коробку, затем осматривает все кругом, включая и коврик, но под ним искать не пытается.

Источник: *The Construction of Reality in the Child* (pp. 86 39–40, 56, 75–76), J. Piaget, 1954, New York, Basic Books, Inc.

же, к примеру, может накинуть на игрушку носовой платок, или зажать в руке маленькую игрушку и перепрятать ее. Здесь заметна характерная особенность подхода Пиаже, о которой говорилось выше: сочетание спонтанного поведения с экспериментально вызванным.

Основной вопрос, который встает после исчезновения предмета: осознает ли младенец то, что предмет все еще существует. Поскольку младенец не умеет говорить, нужно решить, по каким наблюдаемым действиям мы могли бы судить о знаниях младенца. Пиаже обращает внимание преимущественно на поисковые действия. Пытается ли малыш следить за объектом, удаляющимся из поля зрения? Высовывается ли малыш из кроватки и смотрит ли на пол, если выпала игрушка? Пытается ли он снять накинутую на игрушку простыню? Поворачивается ли он незамедлительно к тому месту, где оставил игрушку, даже если некоторое время он на нее не смотрел? Существует множество форм поведения, свидетельствующего о том, что младенец знает или не знает о продолжении существования объекта. Для подхода Пиаже является характерным изучение широкого диапазона соответствующих форм поведения. Для него также характерно внимание к такому поведению, особенно на более высоких уровнях развития, которое связано с активной двигательной реакцией со стороны малыша.

Так же, как и поисковое поведение, исчезновение объекта может принимать разнообразные формы. Привели к исчезновению объекта действия самого младенца или неких внешних сил? Исчез объект целиком, или некоторая его часть все же видна? Если объект исчез целиком, остались ли какие-либо звуковые или тактильные признаки его присутствия? Если используется несколько мест, в которых прячется объект, есть ли у ребенка возможность узнать не только о том, что объект продолжает свое существование, но и о том, *где* он находится? За этими особенностями скрывается еще одна характерная особенность подхода Пиаже: использование при изучении развития любого понятия множества различных задач. Из этой методологии с необходимостью следует теоретический принцип: положение о том, что важнейшие когнитивные навыки, такие как понятие об объекте, не появляются в своем зрелом состоянии, а постепенно и поэтапно развиваются. Именно с тем, чтобы отследить переход от одной стадии к другой, и использует Пиаже множество задач и множество форм поведения. В вопросах, которые остались неразрешенными в изолированных исследованиях Пиаже, нет недостатка. Дальнейшие исследования в этом направлении принимают разнообразные формы, однако есть нечто, что их объединяет. Практически все современные исследования проводятся на больших и более репрезентативных выборках, чем выборка Пиаже. Наряду с увеличением объема выборки уделяется внимание стандартизации и экспериментальному контролю, что редко можно обнаружить в работах Пиаже. Проведение исследований когнитивной сферы в младенчестве в лаборатории стало нормой; исследования в домашней обстановке проводятся очень редко, еще реже исследования бывают по-настоящему натуралистичными. Лабораторные условия дали отсутствовавшую у Пиаже возможность использовать разнообразные технические средства (например, видеозапись). В целом, исследователи пытаются сегодня отвлечься от некоего своеобразия подхода Пиаже, сохранив при этом ориентированность на базовые, интересные с точки зрения психологии развития формы знания.

На специфическом уровне проверка положений теории Пиаже, касающихся развития в период младенчества, проводится в нескольких направлениях. Часто анализируемый и принципиальный вопрос имеет отношение к адекватности использовавшихся Пиаже методик оценки. Действительно ли наблюдения и простые эксперименты Пиаже дают возможность получить точную картину того, что понимает или не понимает ребенок? Или есть случаи, когда Пиаже дает неправильную оценку возможностям малыша? В частности, есть ли ситуации, когда Пиаже недооценивает способности ребенка. Возможно, в основном по причине чрезмерного внимания, которое он уделяет наблюдаемым действиям? Представляется правдоподобным, к примеру, что младенец отлично понимает, что объект продолжает свое существование, однако, вопреки ожиданиям Пиаже, не в состоянии предпринять активных действий по его поиску.

Как еще можно было бы изучить представления младенца об объектах? Наиболее информативен подход, основанный на анализе реакций малыша в случаях, когда принципы постоянства объекта так или иначе нарушаются. Предположим, что объект волшебным образом растворяется в воздухе, или что один твердый предмет проходит через другой. *Вас* бы это, несомненно, озадачило или удивило. Вероятно, и младенца тоже — и это свидетельствовало бы о наличии элементарных представлений о постоянстве объекта даже при отсутствии поискового поведения.

В качестве примера возьмем исследование Байаржон (Baillargeon, 1987). В исследовании Байаржон использовалась описанная ранее в этой главе процедура габитуации/дегабитуации. Сначала младенцы наблюдали за установкой, изображенной в верхней части рис. 10.4 — то есть экран, как подъемный мост, описывающий одним концом дугу в 180° . Хотя первоначально это зрелище вызвало определенный интерес, в конечном итоге малыши стали смотреть на эту установку все меньше и меньше, что являлось свидетельством процесса габитуации, обусловленного многократным предъявлением стимула. В этот момент на пути экрана ставили деревянную коробку (см. нижнюю часть рисунка). Как явствует из рисунка, коробка была видна с начала проверки, но затем закрывалась вставившим перпендикулярно экраном. Сравнивались два экспериментальных условия. При условии «Возможное событие» экран достигал точки соприкосновения с коробкой и останавливался — как должно быть в ситуации, когда на его пути стоит твердый объект. При условии «Невозможное событие» экран доходил до точки соприкосновения с коробкой, а затем (благодаря скрытой платформе, опускавшей коробку) продолжал двигаться! Вопрос состоял в том, будут ли младенцы по-разному реагировать на эти события — то есть вызовет ли одно событие более интенсивную дегабитуацию, чем другое? Ответ оказался утвердительным, так же, как если бы дело касалось вас или меня. Когда экран при соприкосновении с коробкой останавливался, интерес оставался невысоким; однако когда все выглядело так, будто экран проходит через место, занимаемое коробкой, время, которое дети смотрели на происходящее, существенно увеличивалось. Наиболее правдоподобное объяснение этому заключается в том, что младенцы осознавали тот факт, что коробка должна все еще существовать, и поэтому были удивлены прохождением через нее другого предмета. Метод «подъемного моста» — лишь один из ряда изохренных процедур, при помощи которых Байаржон и другие исследователи — наиболее известными из

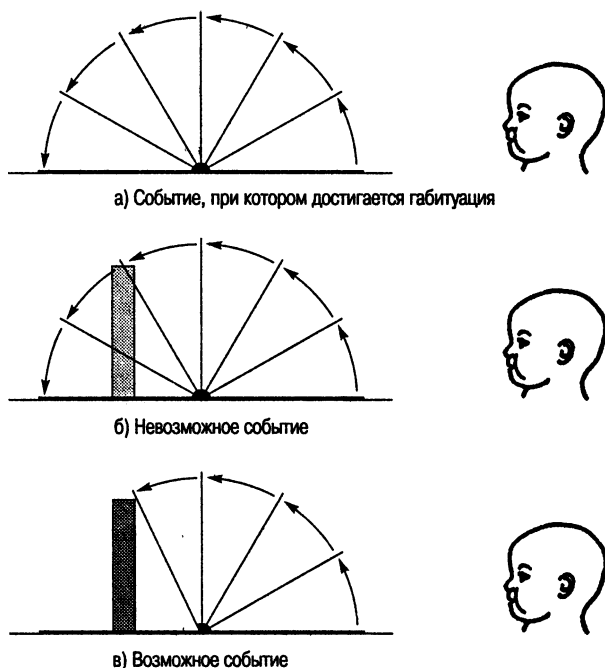


Рис. 10.4. Разработанный Байаржон тест постоянства объекта.

Сначала достигается габитуация на событие (а), после чего измеряется реакция на событие (б) или на событие (в).
(Источник: «Object Permanence in 3,5- and 4,5-Month-Old Infants», R. Baillargeon, 1987, *Developmental Psychology*, 23, p. 656)

которых являются Элизабет Спелке с коллегами — изучают знания малышей о физическом мире (Baillargeon, 1993, 1994; Spelke, 1991). Исследовались разнообразные формы понимания, включая понимание законов тяготения и опоры, принципов причинности, различия между одушевленными и неодушевленными предметами, а также уровни постоянства объекта выше знаний, соответствующих «Стадии 4», выделенной Байаржон (Baillargeon, 1987). Эту работу отличает две методологические особенности. Первая — использование парадигмы несоответствия ожиданиям — создание разнообразных экспериментальных ситуаций, в которых внешне нарушаются некие физические принципы (например, принцип постоянства объектов, закон тяготения). Другая особенность — измерение реакций, требующих от ребенка минимальных моторных усилий, — чаще всего измерение времени фиксации взгляда, но иногда и других параметров (например, изменения частоты сердечных сокращений, выражения лица). И хотя здесь нас интересует больше методология, а не результаты, стоит отметить, что общий вывод всех этих исследований таков: младенцы действительно знают о физическом мире значительно больше, чем это следует из работ Пиаже.

Социальное развитие в младенчестве

Так же, как перцептивная и когнитивная сфера, социальное развитие в младенчестве — большая тема. Посвященный ей раздел делится на две части. Я начну с некоторых общих замечаний о том, как обычно исследуется социальный мир младенца, а затем перейду к двум, наверное, самым популярным вопросам исследований, касающихся социального развития: к понятиям *привязанности* и *темперамента*.

Опять-таки для внесения ясности имеет смысл сравнить новую тему с темами, уже рассмотренными в данной главе. Здесь есть и сходства, и различия. С одной стороны, перед исследователем, изучающим социальное развитие в младенчестве, встают те же проблемы, что и перед исследователем, изучающим перцептивную или когнитивную сферу младенца. В частности, младенец остается бессловесным, зачастую строптивым испытуемым, которого, во-первых, трудно привлечь к исследованию, а во-вторых, трудно удержать в исследовании. С другой стороны, некоторые формы «проблемного поведения», доставляющие множество хлопот тем, кто занимается вопросами познавательной сферы, уже не являются «проблемными», поскольку представляют интерес сами по себе. Начинает ли ребенок плакать, когда мать собирается оставить его одного в экспериментальных условиях? Обращает ли ребенок внимание исключительно на мать, отказываясь реагировать на стимулы, предлагаемые незнакомым взрослым? Такое поведение, являющееся досадной помехой в одних исследованиях, может представлять собой именно то, что намеревается изучить специалист, занимающийся социальным развитием.

Есть и другое отличие. Как отмечалось, исследования, направленные на изучение восприятия в младенчестве, так же, как и большинство постпиажетианских исследований когнитивной сферы младенца, проводятся в лабораторных условиях. Доля исследований социального развития, проводимых в естественной обстановке, то есть дома у ребенка, значительно выше. Даже когда лабораторные условия *используются*, обстановка лаборатории имеет намного больше общего с естественными условиями, чем обстановка лаборатории в любых других исследованиях младенцев. В экспериментальной комнате можно скорее найти игрушки и книжки-раскраски, а не тахистоскопы и электрокардиографы. Да, это все еще незнакомая комната, и многие младенцы совершенно четко это осознают. Действительно, один из достоверных фактов, выявленных в исследованиях социального развития в младенчестве, состоит в том, что в лаборатории разнообразные реакции дистресса — например, плач при разлучении с матерью, страх при виде незнакомого человека — выражены в большей степени, чем в домашней обстановке. Вскоре мы еще вернемся к этому моменту.

Можно указать еще два отличия исследований социального развития от исследований, направленных на изучение перцептивной и когнитивной сферы. Первое из них заключается в термине «социальное». Здесь нас интересует не ребенок, предоставленный самому себе, а ребенок, взаимодействующий с другими людьми. Чаще всего в качестве другого человека выступает мать. В одних случаях в своем поведении мать руководствуются инструкциями экспериментатора; в других — ведет себя свободно, и в действительности ее поведение становится одной из изучаемых

переменных. Хотя традиционно основное внимание уделяется матери, в последние годы в некоторых исследовательских программах отцам также отводится довольно существенная роль. Иногда изучается взаимодействие с другими детьми, особенно по достижении ребенком 18–24-месячного возраста. А поведение с этими знакомыми людьми часто сопоставляется с реакциями на «незнакомцев», то есть тех, с кем ребенок до этого не сталкивался.

Второе отличие касается средств измерения. Изучение восприятия основано преимущественно на автоматической регистрации реакций. Хотя Пиаже проводил свои исследования методом наблюдения, в современных работах, посвященных когнитивной сфере младенца, также наблюдается тенденция к использованию автоматической регистрации — к примеру, движения глаз, частоты сердечных сокращений. Исследования социального развития в большинстве своем основаны на оценках наблюдателей. Для облегчения работы наблюдателей нередко используются такие сложные технические средства, как видеокамеры и регистраторы событий; тем не менее то, что будет зафиксировано, зависит в конечном счете от мнения наблюдателя. Эта опора исключительно на наблюдательную оценку с необходимостью вытекает из цели исследований социального развития: выявить естественные, социально значимые единицы реагирования — «улыбка», «плач», «поиск контакта», «сопротивление». Технические средства могут в этом помочь, однако необходимым звеном остается наблюдатель.

Привязанность

Обратимся теперь к рассмотрению исследований **привязанности**. Привязанность — это глобальный конструкт, заключающий в себе множество феноменов в социальном развитии младенца. Суть этого понятия сводится к тому, что иногда называют «первыми отношениями любви»: к тесной эмоциональной связи, формирующейся между младенцем и его родителями. Привязанность — это глобальный конструкт, заключающий в себе множество феноменов в социальном развитии младенца. Суть этого понятия сводится к тому, что иногда называют «первыми отношениями любви»: к тесной эмоциональной связи, формирующейся между младенцем и его родителями. Нас интересует, как формируется эта связь, а также почему у одних малышей привязанность менее стойкая, чем у других. Наверное, нет более важного вопроса в психологии развития, чем этот.

С точки зрения методологии, изучение привязанности ставит перед исследователем несколько трудностей. Первый вопрос касается поведения, по которому можно судить о наличии или об отсутствии привязанности. «Привязанность», как и «понятие об объекте», «восприятие глубины» или любой другой конструкт, изучаемый в психологии развития, не некая поведенческая данность; это конструкт высокого уровня, о котором приходится судить по разнообразным поведенческим реакциям. Последние 30 лет идет постоянное расширение диапазона реакций, которые учитываются при оценке привязанности. В первых систематических исследованиях этого конструкта (например, Schaffer & Emerson, 1964) основное внимание уделялось *дистрессу при разлучении* — то есть огорчению младенца при разлучении с объектом привязанности. Например, плачет ли малыш, когда мать выходит из комнаты или когда на ночь его одного оставляют в кроватке? Дистресс при раз-

лучении интуитивно воспринимается как проявление привязанности; от малыша, сформировавшего эмоциональную связь с другим человеком, мы ожидаем, что он будет противиться разлучению, в отличие от малыша, который еще не сформировал привязанность. Кроме того, дистресс при разлучении проходит в своем развитии определенные этапы, которые можно ожидать от проявления привязанности; в первые месяцы жизни он отсутствует, однако наблюдается, зачастую в довольно интенсивной форме, в период между 6 и 12 месяцами жизни.

В современных исследованиях сохраняется интерес к дистрессу при разлучении как к одному из важнейших параметров, однако при оценке привязанности к нему добавляется еще ряд поведенческих проявлений. К этим дополнительным реакциям относят позитивные способы проявления младенцем своей привязанности к родителю. Улыбается ли, гулит ли малыш в присутствии матери чаще, чем в присутствии других людей? Такая *дифференцированная отзывчивость* не является привязанностью в полном смысле этого слова, однако обычно рассматривается как один из этапов развития зрелой привязанности. Чувствует ли малыш себя увереннее в присутствии матери, легче ли ему решиться на какой-то шаг, исследовать нечто новое для него, взаимодействовать с незнакомыми людьми? На способность использовать мать как *надежную опору* впервые обратил внимание Харлоу (Harlow, 1958), работавший с детенышами обезьян; оказалось, что она важна и для людей. Радует ли малыша, когда мать заходит в комнату, наблюдается ли у него *поведение приветствия* или *поиск контакта*? В целом, создается ли впечатление, что малыш доволен присутствием матери и ведет себя так, чтобы либо удержать ее рядом с собой, если она уже пришла, либо вернуть ее, если она уходит? Форма, которую принимает поведение привязанности, зависит от ситуации, уровня развития ребенка и его индивидуальных особенностей; однако все виды поведения привязанности объединяет одно — удовольствие и уверенность в присутствии объекта привязанности.

Выбор для изучения конкретных поведенческих проявлений привязанности — лишь один из этапов процесса исследования. Следующий важный этап — выбор способа получения информации об этих поведенческих проявлениях. Как узнать, сопротивляется ли малыш разлучению с матерью, реагирует ли он более позитивно на знакомых людей, использует ли мать как надежную опору и делает ли вообще что бы то ни было, что, по нашему мнению, свидетельствует о наличии привязанности? Ответ на этот вопрос состоит в том, что существует три направления изучения этой проблемы. Можно расспросить кого-либо, кто знает ребенка, как правило, его родителей (метод *родительского отчета*); можно наблюдать поведение ребенка в домашней обстановке (метод *естественного наблюдения*); либо можно создать лабораторную ситуацию, провоцирующую интересующее нас поведение (метод *структурированных лабораторных условий*). Более подробно эти подходы будут обсуждены в главе 12, при рассмотрении социального развития в целом. Здесь же речь пойдет об их применении к явлению привязанности. Начнем с подхода, господствующего в подобных исследованиях все последние годы: использование для оценки привязанности стандартизированных лабораторных условий.

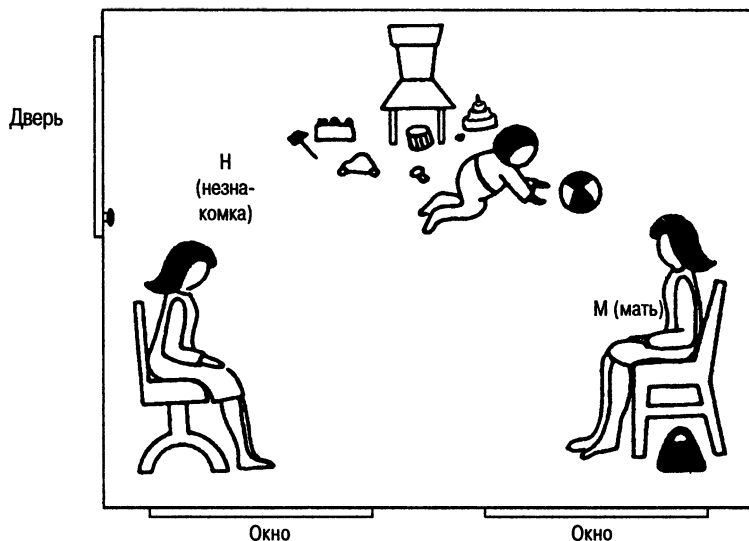


Рис. 10.5. Обстановка, создаваемая для оценки привязанности в «Незнакомой ситуации».
(Источник: «Responses of One-Year-Olds to a Stranger in a Strange Situation» (p. 134),
I. Bretherton & M. D. S. Ainsworth. In M. Lewis & L. A. Rosenblum (Eds.),
The Origins of Fear (pp. 131–164), 1974, New York: Wiley)

Мэри Эйнсворт с коллегами создала наиболее эффективную методику измерения привязанности. Эйнсворт, Блехар, Уотерс и Уолл (Ainsworth, Blehar, Waters, & Wall, 1978) разработали процедуру, которую назвали **«Незнакомой ситуацией»**. Обстановка, создаваемая для «Незнакомой ситуации», изображена на рис. 10.5. Как явствует из рисунка, в ситуации принимает участие три человека: мать, незнакомка и малыш. Помещение находится в университетском здании и не знакомо ни малышу, ни матери; поэтому обстановку называют лабораторной, а не естественной. Однако в этом помещении создаются комфортные условия, даже несколько напоминающие условия игровой комнаты: стулья и журналы для взрослых, яркие картинки на стенах и множество игрушек для малыша. Хотя на рисунке этого не видно, есть еще одна особенность: двустороннее зеркало, через которое, оставаясь незаметным, можно следить за происходящим в комнате. Мать знает об этом зеркале и о находящемся за ним наблюдателе, а ребенок — нет.

«Незнакомую ситуацию» составляет последовательность из восьми эпизодов, представленная в табл. 10.3. Во всех восьми эпизодах поведение матери по большей части (хотя и не целиком) контролируется. Поведение же малыша не контролируется, и именно малыш является объектом изучения. Как следует из приведенного в таблице описания, эпизоды спланированы таким образом, чтобы вызвать разнообразные формы поведения привязанности. Наблюдается взаимодействие малыша с матерью до и после разлучения с ней, а также поведение в одиночестве и в присутствии незнакомки. Кроме того, наблюдается поведение малыша с незнакомкой, которое сравнивается с его поведением в присутствии матери. Последовательность событий включает как эпизоды «разлучения», так и эпизоды «воссоединения»; в отсутствие матери ребенок остается либо один, либо с незнакомкой. В целом, проце-

дура являет собой попытку уложить в 20-минутный интервал множество ситуаций и форм поведения, используемых для изучения привязанности.

Таблица 10.3

Эпизоды «Незнакомой ситуации»

Номер эпизода	Участники	Продолжительность	Краткое описание действия
1	Мать, малыш и наблюдатель	30 с	Наблюдатель вводит мать и малыша в экспериментальную комнату и покидает их
2	Мать и малыш	3 мин	Мать ничего не предпринимает, пока малыш исследует комнату; по прошествии 2 минут инициируется игра
3	Незнакомка, мать и малыш	3 мин	Входит незнакомка. Первая минута: незнакомка молчит Вторая минута: незнакомка разговаривает с матерью Третья минута: незнакомка подходит к малышу. По прошествии 3 минут, мать, стараясь не привлекать внимание малыша, выходит
4	Незнакомка и малыш	3 мин или менее ^а	Первый эпизод разлучения. Незнакомка подстраивает свое поведение под поведение малыша
5	Мать и малыш	3 мин. или более ^б	Первый эпизод воссоединения. Мать приветствует и/или успокаивает малыша, затем пытается вновь привлечь его внимание к игре. Затем со словами прощания мать выходит
6	Малыш в одиночестве	3 мин или менее ^а	Второй эпизод разлучения
7	Незнакомка и малыш	3 мин или менее ^а	Продолжение второго эпизода разлучения. Входит незнакомка и подстраивает свое поведение под поведение малыша
8	Мать и малыш	3 мин	Второй эпизод воссоединения. Входит мать, она приветствует малыша и берет его на руки. Незнакомка, стараясь не привлекать внимание, выходит

^а Длительность эпизода уменьшается, если малыш слишком расстроен.

^б Длительность эпизода увеличивается, если малышу требуется больше времени для возвращения к игре.

Эйнсворт и ее коллеги используют метод непосредственного наблюдения поведения младенца. Оценка производится на трех уровнях. На микроуровне фиксируются специфические, дискретные поведенческие акты. Примерами категорий этого уровня служат плач, улыбка, вокализация, телодвижения. На втором уровне производится оценка взаимодействия: близости или стремления к контакту; поддержания контакта; сопротивления; избегания; поиска; дистантного взаимодействия. Оценка на этом уровне предполагает более свободную интерпретацию, чем оценка на первом уровне (например, уже «сопротивление», а не «плач»); еще одно отличие заключается в том, что категории второго уровня носят явно социальный характер. Наконец, третий уровень оценки связан с общей квалификацией отношения привязанности между младенцем и матерью. В большинстве исследований используются три глобальные категории. Самый адаптивный из трех моделей привязанности — тип *B* (*надежная привязанность*), как обозначают устойчивую и адекватную привязанность. Младенец, относящийся к типу *B*, очевидно чувствует себя наиболее комфортно в присутствии матери и огорчается, когда она уходит; однако он не впадает в отчаяние при отсутствии матери и, в целом, способен достаточно хорошо адаптироваться ко всем трудностям, связанным с нахождением в незнакомой среде, с присутствием незнакомого человека, с приходами и уходами матери. Две другие формы привязанности не столь устойчивы и адекватны. Младенец, относящийся к типу *A* (*избегающей привязанностью*) при разлучении с матерью не проявляет практически никакого огорчения, а по ее возвращении не выказывает особой радости; кроме того, такой младенец ведет себя с незнакомыми людьми почти так же, как с матерью. Младенец, относящийся к типу *C* (*амбивалентной привязанностью*) выглядит более расстроенным и раздражительным. Такого малыша разлучение с матерью может привести в отчаяние, а по возвращении матери он либо активно ее отвергает, либо ведет себя амбивалентно: то тянется к ней, то отталкивает ее.

Недавно некоторые исследователи (например, Main & Solomon, 1991) стали выделять четвертый вид привязанности, тип *D*, или *дезорганизованную привязанность*. Тип *D* — это еще одна форма тревожной привязанности; ее характеризуют смешанные проявления избегающего и амбивалентного поведения; при этом часто все выглядит так, будто воссоединение с матерью приводит младенца в замешательство и вызывает опасения.

Как следует из приведенного выше описания, «Незнакомая ситуация» провоцирует разнообразные поведенческие проявления привязанности. Главное достоинство структурированной лабораторной оценки состоит именно в этом разнообразии; данный подход и экономичнее естественного подхода, и, в отличие от метода родительских отчетов, предполагает непосредственную оценку поведения ребенка. С другой стороны, в «Незнакомой ситуации» теряется преимущество двух других подходов — изучение поведения в естественных условиях. Однако, несмотря на недостатки лабораторной обстановки и прилагательное «незнакомая», опыт, получаемый младенцем при использовании этого метода, как правило, не является чем-то новым. Матери и в реальной жизни приносят малышей в незнакомые комнаты; малыши и в реальной жизни сталкиваются с незнакомыми людьми; их и в реальной жизни иногда оставляют в одиночестве и т. д. Что делает «Незнакомая

ситуация», так это сокращает время, в течение которого ребенок получает весь этот опыт. Это сокращение, равно как и новизна обстановки, несомненно, являются стрессом для младенца; и в действительности, в «Незнакомой ситуации» младенцы демонстрируют дистресс чаще, чем дома. Однако Эйнсворт с соавторами (Ainsworth, Blehar, Waters, & Wall, 1978) утверждают, что накопление стресса делает ситуацию более информативной, позволяя выявить поведение привязанности. По их словам,

«Именно потому, что эта процедура своим кумулятивным характером с нарастающей интенсивностью провоцирует поведение привязанности, появляется возможность в сравнительно короткий промежуток времени наблюдать поведение привязанности в условиях перехода от сравнительно слабой к весьма сильной активации. В знакомой домашней среде редко происходит сильная активация поведения привязанности, поэтому для выявления такого же диапазона реакций потребовались бы долгие часы наблюдения...» (р. xi).

Даже если мы примем этот аргумент, остается необходимость удостовериться в том, что такое же проявление привязанности имеет место и в домашней обстановке. Как уже отмечалось, сделать это можно двумя способами: расспросить родителей о том, как все происходит обычно, или посетить малыша дома и понаблюдать за его поведением. В работах, посвященных привязанности, встречаются оба подхода. Исследовательская программа самой Эйнсворт, к примеру, включает помимо «Незнакомой ситуации» интервью и наблюдение (Ainsworth, 1967; Ainsworth et al., 1978); в действительности, лабораторная процедура основывается на проведенной ранее работе в естественных условиях. Пример, который мы рассмотрим здесь, относится к одной из последних исследовательских разработок — сегодня это главная альтернатива «Незнакомой ситуации» как метода оценки привязанности. Речь идет о **Q-наборе «Привязанность»** (*Attachment Q-Set* или *AQS*), разработанном Эверетом Уотерсом с коллегами (Waters, 1995; Waters & Deane, 1985).

Q-набор «Привязанность» — частный случай метода оценки индивидуальных различий, который называют *методом q-сортировки*. Его применение к привязанности ничем не отличается от применения его к другим конструктам. *AQS* состоит из 90 утверждений, которыми описывается поведение младенца или особенности, характеризующие отношение привязанности. В табл. 10.4 выборочно представлены отдельные пункты *AQS*. Заметьте, что в одних случаях (например, в первых нескольких пунктах) об адекватной привязанности свидетельствует позитивная реакция, тогда как в других случаях (например, в двух последних пунктах) желательнее именно отсутствие указанных характеристик. Как любая *q*-сортировка, *AQS* предполагает последовательное отнесение утверждений к одной из категорий. Сначала лицо, оценивающее ребенка, приблизительно поровну делит 90 утверждений на три категории: «описывает ребенка», «не описывает ребенка» и «трудно сказать». За этим первичным делением следует разбиение каждой из категорий на три более дифференцированные подкатегории, в каждую из которых относят по 10 утверждений. Таким образом, конечный результат — 9-уровневое ранжирование характеристик младенца, начиная с 10 наиболее типичных и кончая 10 наименее типичными. Именно эта общая картина, основанная на структурированной информации, и представляет собой оценку привязанности ребенка.

Таблица 10.4

Примеры утверждений из Q-набора «Привязанность»

Номер утверждения	Описание
1	Ребенок охотно делится с матерью или дает ей подержать то, что она просит
11	Ребенок часто по собственной инициативе обнимает мать и прижимается к ней
21	Играя во дворе, ребенок следит за тем, где находится мать Время от времени он подзывает ее; следит за ее переходами из комнаты в комнату
71	На руках у матери ребенок прекращает плакать и быстро оправляется от страха или огорчения
73	У ребенка есть игрушка, которую он кладет с собой в кровать и за которую держится, когда бывает расстроен, или одеяльце, в которое он кутается в таких же ситуациях
76	Имея выбор, ребенок предпочитает играть с игрушками, а не с взрослым
79	Ребенок быстро раздражается на мать
81	Ребенок использует плач как способ получения от матери того, что он хочет

Источник: «The Attachment Q-Set (Version 3.0)», E. Waters, 1995, *Monographs of the Society in Child Development*, 60 (2–3, Serial No. 244), p. 236–246.

Я уже отмечал, что альтернативой лабораторного исследования является либо наблюдение в домашней обстановке, либо интервью с родителями. *AQS* можно использовать при любом из этих способов сбора данных. В одних случаях специально подготовленные наблюдатели в течение нескольких часов следят за взаимодействием матери и младенца в домашней обстановке (например, Strayer, Verissimo, Vaughn, & Howes, 1995), в других случаях матери сами классифицируют утверждения *AQS* (например, Posada et al., 1995). Как правило, рейтинговые оценки, которые, используя *AQS*, дают независимо друг от друга матери и наблюдатели, очень близки (Waters & Deane, 1985). Оценки *AQS* также в значительной мере коррелируют с результатами разделения на группы при использовании «Незнакомой ситуации» (Vaughn & Waters, 1990).

Назначение «Незнакомой ситуации» и Q-набора «Привязанность» одно — оценка индивидуальных различий в привязанности. Что еще можно было бы обнаружить, зная о наличии таких различий — то есть какие еще виды исследования, помимо этой базовой оценки, интересно было бы провести? Очевидно, огромное значение имеют ответы на два вопроса. Первый касается *причин* — чем обусловлены эти различия? Исследователи привязанности анализируют причины разными способами. Наверное, чаще всего, начиная с некоторых исследований самой Эйнсворт, предпринимается попытка выявить используемые матерью *методы воспитания*, лежащие в основе разных моделей привязанности. Методы воспитания одновременно и наиболее важная и наиболее трудная с методологической точки зрения

область исследования; детально мы рассмотрим ее в главе 12. Еще один вариант — *кросс-культурные исследования* — анализ того, в какой мере различаются модели привязанности в разных культурных условиях, особенно в условиях, где опыт младенца несколько отличается от опыта младенцев западных культур, в которых чаще всего проводятся исследования. Иногда определенные культуральные различия действительно обнаруживаются, хотя их интерпретация все еще остается спорным вопросом (Posada et al., 1995; Van IJendorn & Kroonenberg, 1988). Наконец, помимо социальных истоков привязанности, исследователи пытаются обнаружить ее биологическую основу. Чаще всего изучается биологический фактор, который мы рассмотрим в следующем разделе: врожденные различия в *темпераменте*. Результаты исследований свидетельствуют о том, что темперамент действительно связан с типом привязанности, хотя здесь тоже есть определенные противоречия в отношении того, насколько сильна эта связь и каковы ее причины (Goldsmith & Alansky, 1987; Vaughn et al., 1992).

Помимо причин нам, конечно, хотелось бы знать о *следствиях* отношений привязанности в раннем детстве. Отражается ли качество привязанности в младенчестве на дальнейшем развитии ребенка? Чтобы ответить на этот вопрос, нужно использовать лонгитюдный план исследования, при котором наблюдалось бы развитие одних и тех же детей, начиная с младенчества и кончая каким-то моментом позднего детства. На сегодняшний момент проведено множество таких исследований, и в них были обнаружены разнообразные связи между привязанностью в раннем детстве и последующим социальным и когнитивным функционированием. Обнаружилось, например, что дети, для которых характерна надежная привязанность в младенчестве, в дошкольном возрасте в среднем по выборке более компетентны в социальном плане и имеют более высокую самооценку, чем дети с менее надежной привязанностью (Sroufe, 1983). Аналогично, знание типа привязанности в младенчестве позволяет с определенной долей уверенности спрогнозировать успешность выполнения когнитивных тестов в дошкольном возрасте и популярность среди сверстников — в младшем школьном возрасте (Arend, Gove, & Sroufe, 1979; Elicker, Englund, & Sroufe, 1992). Как и любой прогноз на основе данных, полученный при обследовании младенца, прогноз, устанавливающий зависимость между привязанностью в раннем детстве и какой-либо личностной характеристикой на дальнейших этапах развития, далеко не точен. Тем не менее результаты практических исследований подтверждают мнение теоретиков: адекватная привязанность в раннем возрасте направляет развитие на оптимальный путь.

Темперамент

Авторами первых работ, посвященных темпераменту в детском возрасте, стали Томас, Чесс и Берч (Thomas & Chess, 1977; Thomas, Chess, & Birch, 1968), исследования которых проводились в рамках нью-йоркского лонгитюдного исследования (New York Longitudinal Study). Начнем поэтому с цитаты из работы этих авторов, обобщающей то, что они понимают под термином «темперамент»:

«Темперамент следует рассматривать в качестве основного термина, характеризующего поведение с точки зрения *«каким образом»*. Он отличается от способностей, которые характеризуют поведение с точки зрения *«что»* и *«насколько хорошо»*; он отличается и

от мотивации, объясняющей, «почему» человек делает то, что он делает. Темперамент характеризует *манеру* поведения человека. ...Понятие темперамента по значению можно приравнять к понятию *стиль поведения*». (Thomas & Chess, 1977, p. 9).

Объектом целенаправленного научного изучения темперамент стал сравнительно недавно. Однако большинство идей, лежащих в основе подобных исследований, далеко не новы, поскольку отражают взгляды, которых придерживаются многие родители. Дети — даже те из них, кто воспитывается в одной семье, — часто совершенно по-разному осваивают окружающий мир. Одни дети полны энергии и постоянно находятся в движении; другие не столь подвижны и предпочитают спокойные занятия. Одни дети могут довольно долго удерживать внимание и доводят свои дела до конца; другие — быстро отвлекаются и берутся то за одно, то за другое. Стиль поведения может не изменяться от ситуации к ситуации и от одного вида деятельности к другому. Кроме того, он может не изменяться с возрастом, впервые проявившись в младенчестве и оставаясь характерной особенностью личности многие годы. Он может проявиться так рано, что будет восприниматься как нечто, по крайней мере, частично обусловленное биологически, а не социально.

Как отмечалось, Томас, Чесс и Берч были первыми, кто начал систематическое изучение темперамента, и их работа послужила образцом для дальнейших исследований и развития теории темперамента. Эти исследователи выделяют девять параметров темперамента, перечисленных и кратко описанных во вставке 10.3. Все параметры оцениваются по трехбалльной шкале, хотя в разных категориях определения каждого из трех делений шкалы различаются. Уровень активности, к примеру, оценивается как «высокий», «средний» и «низкий». Интенсивность реакции оценивается как «позитивная», «изменчивая» и «негативная». Другие концепции темперамента обычно представляют собой вариант многих из этих параметров, хотя могут включать и дополнительные факторы. Басс и Пломин (Buss & Plomin, 1984), к примеру, подчеркивают значение таких параметров, как эмоциональность, активность и общительность. Каган и коллеги (например, Kagan, Reznick, & Gibbons, 1989) акцентируют внимание на раннем проявлении модели заторможенного поведения.

Как происходит сбор данных в исследованиях Томаса с соавторами? Ответ, который имеет отношение не только к работам Томаса, но и к большинству исследований, посвященных темпераменту, заключается в том, что исследователи узнают о темпераменте ребенка, расспрашивая о нем его родителей. В нью-йоркском лонгитюдном исследовании основной блок данных был получен при помощи интервью с родителями. Первые интервью проводились тогда, когда дети находились на 2–6 месяце жизни. В дальнейшем интервью проводились каждые 3 месяца в течение первого года жизни ребенка и с интервалом в 6 месяцев — на втором году жизни. В интервью речь шла о разнообразных повседневных ситуациях, в которых может проявиться темперамент ребенка. Родителей спрашивали, к примеру, о том, как ребенок спит и ест; о типичном для него поведении при встрече с незнакомым человеком или на приеме у врача; о том, как он реагирует, когда его моют, переодевают, меняют ему пеленки. Для повышения точности отчетов была предпринята попытка увязать эти вопросы с конкретными жизненными ситуациями и поведением, которое наблюдалось совсем недавно. Кроме того, исследователи стремились

Вставка 10.3

Параметры темперамента, выделенные в нью-йоркском лонгитюдном исследовании

Уровень активности: моторный компонент, проявляющийся в деятельности ребенка, а также в суточном соотношении периодов активности и пассивности. При оценке этой категории используются протокольные данные о подвижности во время купания, еды, игры, одевания и других процедур ухода, а также информация, касающаяся цикла «сон—бодрствование», реакции хватания, ползания и ходьбы.

Ритмичность (регулярность): предсказуемость и/или непредсказуемость времени любого вида деятельности. Может анализироваться в связи с циклом «сон—бодрствование», требованием пищи, режимом кормления и очищения организма.

Приближение/уход: характер первой реакции на новый стимул, будь то новая пища, новая игрушка или новый человек. Реакции приближения позитивны, проявляются ли они в эмоциях (улыбка, вербализации и т. д.) или в действиях (прием новой пищи, попытка схватить новую игрушку, активная игра и т. д.). Реакции ухода негативны, проявляются ли они в эмоциях (плач, нервозность, гримасы, вербализации и т. д.) или в действиях (избегание, выплевывание новой пищи, отталкивание новой игрушки и т. д.).

Адаптивность: реагирование на новую ситуацию или изменение ситуации. Важен не характер первоначальной реакции, а легкость, с которой она видоизменяется в нужном направлении.

Порог реагирования: уровень интенсивности стимуляции, необходимый, чтобы вызвать какую-либо реакцию, вне зависимости от формы, которую принимает эта реакция. При оценке этого параметра используются реакции на сенсорную стимуляцию, предметы обстановки и социальные контакты.

Интенсивность реакции: сила реакции, вне зависимости от ее качества и направления.

Настроение: соотношение поведенческих проявлений радости, удовольствия, дружеского настроения к поведенческим проявлениям неудовольствия, недружелюбия и плача.

Отвлекаемость: эффективность внешних стимулов в прерывании или изменении направления текущей деятельности ребенка.

Длительность и устойчивость внимания: две взаимосвязанные категории. Длительность внимания характеризуется продолжительностью периода, который ребенок занят некой деятельностью. Устойчивость внимания характеризует способность продолжать деятельность при наличии факторов, препятствующих поддержанию выбранного направления деятельности.

Источник: *Temperament and Behavior* (p. 21–22), A. Thomas & S. Chess. New York: Brunner/Mazel, 1977.

получить описание поведения, а не интерпретацию его родителями (например, «Он выплевывает кашу», а не «Он ненавидит кашу»).

Интервью — это лишь один из методов получения родительских отчетов. Еще один подход — использование опросников, при котором устные ответы интервью заменяют письменными. Формы письменных ответов разнообразны; наиболее простая и распространенная — предложение испытуемому сообщить, в какой мере он согласен с тем или иным утверждением, указав на одно из делений упорядоченной шкалы ответов. В табл. 10.5 приведены примеры утверждений из опросника, предназначенного для изучения младенцев в соответствии с выделенными в исследованиях Томаса и коллег параметрами. Учитывая большую простоту и экономичность опросников, неудивительно, что при изучении темперамента они используются чаще, чем интервью. В одном обзоре приводится 29 методик измерения темперамента, 22 из которых предполагают использование опросника (Hubert, Wachs,

Таблица 10.5

Примеры утверждений из опросника «Темперамент младенца»

Почти никогда	Редко	По-разному, обычно нет	По-разному, обычно да	Часто	Почти всегда
1	2	3	4	5	6
1. Младенец ежедневно ест примерно одинаковое количество твердой пищи			Почти никогда	1 2 3 4 5 6	Почти всегда
2. Перед сном и сразу после просыпания младенец перевозен (хмурится, плачет)			Почти никогда	1 2 3 4 5 6	Почти всегда
3. Младенец играет с одной игрушкой не более минуты, а затем ищет другую игрушку или другое занятие			Почти никогда	1 2 3 4 5 6	Почти всегда
4. При просмотре телевизора или занимаясь похожим видом деятельности, младенец сидит тихо			Почти никогда	1 2 3 4 5 6	Почти всегда
5. Младенец быстро замечает любые перемены места или положения тела при кормлении, а также перемены в человеке, кормящем его			Почти никогда	1 2 3 4 5 6	Почти всегда
6. Младенец без капризов дает подстригать себе ногти			Почти никогда	1 2 3 4 5 6	Почти всегда
7. Младенца, который плачет, требуя пищи, можно успокоить на время более одной минуты, взяв его на руки, дав соску, повязав ему нагрудник и т. п.			Почти никогда	1 2 3 4 5 6	Почти всегда
...					
89. Младенец спокойно переносит купание. Удовольствие или неудовольствие выражаются в мягкой форме (улыбка или хмурое лицо)			Почти никогда	1 2 3 4 5 6	Почти всегда
90. Для введения в рацион ребенка новых продуктов требуется 3 или более попыток			Почти никогда	1 2 3 4 5 6	Почти всегда
91. Первая реакция младенца на любую новую процедуру (первая стрижка, новое лекарство и т. п.) — отказ			Почти никогда	1 2 3 4 5 6	Почти всегда
92. С мокрыми подгузниками ребенок ведет себя так же, как с сухими (отсутствие реакции)			Почти никогда	1 2 3 4 5 6	Почти всегда
93. Младенец капризничает во время медицинского осмотра			Почти никогда	1 2 3 4 5 6	Почти всегда
94. Младенец принимает изменения в рационе и режиме кормления с первой или второй попытки			Почти никогда	1 2 3 4 5 6	Почти всегда
95. Младенец подвижен и может, играя в одиночестве, несколько минут или более двигать ножками и ручками, прыгать			Почти никогда	1 2 3 4 5 6	Почти всегда

Peters-Martin, & Gandour, 1982). По последним оценкам (например, Slabach, Morrow, & Wachs, 1991) опросники остаются наиболее часто используемым методом.

Вне зависимости от того, какая форма исследования используется — интервью или опросники — данные, полученные с помощью этих методов, представляют собой словесные отчеты, а не результаты непосредственного наблюдения. Несмотря на то, что валидность подхода с использованием родительских отчетов является общепризнанной, очевидно, что данные этих методик желательно дополнить данными наблюдения поведения, по которому можно судить о темпераменте. Хотя некоторые из подобных методик разработаны для использования их в домашней обстановке (например, Rothbart, 1980), большинство все же основано на создании определенных ситуаций в структурированных лабораторных условиях. Таблица 10.6 иллюстрирует одну из таких методик, которую разработали Голдсмит и Ротбарт (Goldsmith & Rothbart, 1991, 1992): «Батарей оценки темперамента в лабораторных условиях» (Laboratory Temperament Assessment Battery, или LAB-TAB). Как можно заметить, LAB-TAB представляет собой ряд эпизодов (половина этих эпизодов приведена в таблице), на которые может реагировать младенец, и с помощью которых можно измерить различия по таким параметрам как уровень активности, пугливость и настойчивость.

Таблица 10.6

Примеры ситуаций из «Батарей оценки темперамента в лабораторных условиях»

Эпизоды страха
По столу на ребенка быстро движется игрушечная механическая собака. Незнакомый мужчина подходит к ребенку и берет его на руки.
Эпизоды гнева
Мягкие объятия родителя, сковывающие движения ребенка во время игры. Интересная игрушка, помещенная за барьером из оргстекла.
Эпизоды удовольствия
Реакция на звуки и мигание лампочек в условиях отсутствия социального контакта. Видоизмененная игра в прятки, когда лицо матери то появляется, то исчезает за разными дверцами.
Эпизоды интереса и настойчивости
Ориентация на выполнение при игре с кубиками. Устойчивость внимания при многократном предъявлении фотографических слайдов.
Эпизоды активности
Проявления активности в большом манеже с множеством больших надувных мячей. Локомоторная активность во время свободной игры.

Источник: «Contemporary Instruments for Assessing Temperament by Questionnaire and in the Laboratory» (p. 264), H. H. Goldsmith & M. K. Rothbart. In J. Strehan & A. Angleitner (Eds.), *Explorations in temperament: International perspectives on theory and measurement* (p. 249–272). New York: Plenum Press.

Перейдем теперь от методик к вопросам, на изучение которых они направлены. Перед исследователями темперамента, как и перед исследователями привязанности, стоят два глобальных вопроса: вопрос *причин* и вопрос *следствий*. Что лежит в

основе темперамента и как отражаются особенности темперамента в раннем детстве на дальнейшем развитии ребенка?

Обычно обсуждение причин так или иначе касается вопроса о роли биологических факторов. Во многом интерес к темпераменту связан с перспективами выявления генетических обусловленных различий между детьми, различий, опережающих социализирующее действие родителей и в то же время оказывающих определенное влияние на процесс социализации. В действительности, во многих теориях и определениях понятия «темперамент» ключевым является генетический компонент (например, Buss & Plomin, 1984).

Доказательства того, что темперамент в известной мере определяется генетической основой, разнообразны. Одно из них — тот факт, что индивидуальные различия проявляются уже в самом раннем возрасте, когда вероятность влияния факторов социализации минимальна. В нью-йоркском лонгитюдном исследовании различия были выявлены в группе 2-месячных детей, в самой младшей из обследовавшихся групп. В других исследованиях индивидуальные различия, включая различия по параметрам темперамента, были зафиксированы уже в период новорожденности (Ricciuti & Breitmayer, 1988). Совершенно ясно, что чем раньше в ходе развития проявляются различия, тем более вероятно, что они обусловлены генетически.

Связь генов и темперамента — одна из областей изучения такой науки, как *генетика поведения*. В главе 11, в контексте противопоставления факторов наследственности и среды в отношении индивидуальных различий в уровне интеллекта, мы более подробно остановимся на методах генетики поведения. Замечу лишь, что при изучении темперамента наибольшее значение в генетике поведения имеет *близнецовый метод*, который заключается в сравнении двух типов близнецовых пар, однояйцевых и двуяйцевых. Логика данного подхода проста: если какой-то аспект развития определяется генетическими факторами, однояйцевые близнецы (общность генов которых составляет 100 %) должны больше походить друг на друга, чем двуяйцевые близнецы (общность генов которых составляет только 50 %). Это действительно было подтверждено при изучении темперамента: многие исследователи пришли к единому выводу о том, что по темпераменту однояйцевые близнецы больше похожи друг на друга, чем двуяйцевые (например, Braungart, Plomin, DeFries, & Fulker, 1992; Emde et al., 1992). Большее сходство проявляется уже в 3-месячном возрасте и сохраняется весь период детства; оно характерно для разных параметров темперамента и было обнаружено при помощи разнообразных методик оценки темперамента.

Еще одно замечание, касающееся близнецового метода, подведет нас ко второму глобальному вопросу исследований темперамента. Результаты близнецовых исследований указывают на то, что гены — одна из важнейших детерминант темперамента, но, как обнаружилось, это не единственная причина индивидуальных различий. Однояйцевые близнецы, может быть, и похожи по темпераменту, но они далеко не одинаковы; и любые различия между ними — это, скорее всего, результат воздействия среды. Таким образом, темперамент не является чем-то, полностью предопределенным генами; свою роль играет и среда.

Вывод о значимости среды следует также и из результатов еще одного типа исследований темперамента: исследований его *стабильности* в процессе развития. Остаются ли неизменными индивидуальные различия в темпераменте с возрас-

том? Для получения ответа на этот вопрос необходимо использование лонгитюдного подхода, при котором одних и тех же детей неоднократно обследуют в процессе их развития. Как следует из названия исследовательского проекта Томаса с соавторами, лонгитюдный анализ с самого начала являлся частью исследования темперамента. Первое интервью в нью-йоркском лонгитюдном исследовании проводилось, когда дети находились в младенческом возрасте; затем, при помощи разнообразных методик — опросников для родителей и учителей, наблюдения, интервью с самими детьми — развитие многих испытуемых отслеживалось вплоть до достижения ими подросткового возраста. Общий вывод — который был подтвержден и последними лонгитюдными исследованиями — заключается в том, что для темперамента характерна определенная, хотя и не абсолютная, стабильность. То есть особенности темперамента в раннем детстве действительно отражаются на дальнейшем развитии: застенчивые дети нередко остаются застенчивыми, раздражительные — раздражительными и т. д. Но это лишь тенденция, а не абсолютная неизбежность; процесс развития всегда зависит от взаимодействия биологических факторов и опыта. Если, к примеру, изначально раздражительного и капризного младенца воспитывают терпеливые и отзывчивые родители, трудности раннего детства не обязательно перерастут в проблемы старшего возраста.

Последние годы понятие темперамента стало одной из излюбленных тем для обсуждения, поэтому в дополнительных литературных источниках нет недостатка (Bates & Wachs, 1994; Goldsmith et al., 1987; Halverson, Kohnstamm, & Martin, 1994). Поскольку в этой главе основное внимание уделялось методикам оценки темперамента в младенчестве, должен отметить, что в представленных источниках содержатся методики, разработанные для оценки темперамента в позднем детстве и даже во взрослости (например, McClellan, 1995; Windle & Lerner, 1986). То же замечание справедливо и по отношению к другому конструкту, о котором шла речь в этом разделе, — к привязанности. Хотя привязанность, даже в большей степени, чем темперамент, традиционно считается проблематикой периода младенчества, в последние годы ее изучение распространилось и на период детства, и даже взрослости (например, Oppenheim & Waters, 1995; van IJzendoorn, 1995).

Резюме

Эта глава начинается с рассмотрения некоторых трудностей, связанных с проведением исследования младенцев. Обсуждаются три основные проблемы. Первая — поиск испытуемых-малышей и удержание их в исследовании. Привлечение младенцев может оказаться затруднительным; они составляют группу испытуемых с самым высоким уровнем выбывания. Как следствие, выборка младенцев зачастую не соответствует критерию случайного отбора. Вторая проблема связана с высоким уровнем выбывания; она заключается во влиянии состояния младенца на его поведение. Уровень активности младенца — это важнейший фактор, определяющий его способ реагирования на окружающий мир; более того, часто состояние младенца (например, состояние сонливости или дистресса) исключает возможность оптимального реагирования. Последняя проблема касается методик оценки поведения. Младенцы не умеют говорить, а самые маленькие из них имеют весьма огра-

ниченные моторные навыки; следствием этого является необходимость проявлять известную изобретательность при поиске реакций, по которым можно было бы судить о том, как малыш воспринимает окружающий его мир. Описываются три системы реакций, которые оказались особенно информативными: *зрительные фиксации, сосание и физиологические реакции*.

Затем речь идет о возрастных сравнениях. То, насколько легко можно произвести валидное возрастное сравнение, зависит от возрастного диапазона выборки. Ограничив сферу интересов периодом младенчества, исследователи могут избежать некоторых проблем, характерных для исследований методом поперечных срезов и лонгитюдных исследований (например, трудности разграничения эффектов возраста и других параметров). С другой стороны, при попытках сравнить младенцев с детьми старшего возраста эти проблемы стоят достаточно остро. Особое значение при сравнении любых возрастных групп имеет проблема *эквивалентности измерений*: то есть проблема поиска реакций, психологически равноценных для разных возрастов.

Первая специфическая тема, которая затрагивается в данной главе, — восприятие в младенчестве. При изучении восприятия проблема измерения реакций особенно сложна, поскольку восприятие — это внутриспихический феномен, не всегда имеющий свои четкие поведенческие проявления. Исследования восприятия младенцев направлены в основном на оценку реакций, которые поддаются точному, как правило, автоматическому, фиксированию в лабораторных условиях; еще одно достоинство лабораторных условий — возможность строгого контроля стимулов. Рассматриваются четыре метода изучения восприятия. Метод *предпочтений* используется для изучения вопросов зрительного различения; предъявляются два зрительных стимула, и на основании того, что взгляд младенца фиксируется на одном из стимулов дольше, чем на другом, делается вывод о способности ребенка различать эти стимулы. Метод *габитуации/дегабитуации* имеет более широкие возможности, поскольку применим для изучения любой модальности восприятия. Сначала достигается габитуация на один стимул; возобновление реакции на другой стимул свидетельствует о различении. Для изучения восприятия в младенчестве часто используются разнообразные формы *обусловливания*, включая пресыщение, когда возобновление эффективности подкрепления является признаком различения стимулов, и генерализацию, когда о перцептивном сходстве судят по переносу реакции с одного стимула на другой. Наконец, «*Визуальный обрыв*» используется для изучения одного-единственного вопроса, представляющего значительный теоретический интерес: восприятие младенцем глубины.

Господствующее положение в области изучения когнитивной сферы младенца занимают работы Пиаже. Исследования Пиаже во многом отличаются от исследований восприятия младенцев; в основном тем, что Пиаже делал акцент на наблюдении в естественных условиях и на гибком, ориентированном на открытие зондировании возможностей ребенка. Среди множества элементарных представлений, которые изучал Пиаже, — *понятие об объекте*: понимание того, что объект существует вне зависимости от непосредственного перцептивного контакта. В методах изучения понятия об объекте, которые использовал Пиаже, основной упор делался на поиске младенцем исчезнувших предметов — на разнообразных формах по-

иска и на разнообразных способах исчезновения. Понятие об объекте — одна из наиболее популярных тем постпиажетианских исследований когнитивной сферы младенца. В работах последователей Пиаже уровень стандартизации и контроля выше, а лабораторные условия и автоматизированная фиксация результатов используется чаще. Особый интерес представляют результаты применения методов оценки возможностей ребенка, которые не предъявляют высоких требований к его моторике. Весьма информативным оказался метод измерения времени фиксации взгляда в ответ на явное нарушение физических законов.

Исследования социального развития в меньшей степени ограничены лабораторными условиями, чем большинство исследований, посвященных перцептивной и познавательной сфере. Здесь в центре внимания находится взаимодействие младенца с другими людьми, чаще всего с матерью. Особый интерес представляет развитие *привязанности* — эмоциональной связи с заботящимся о ребенке лицом, имеющей разнообразные поведенческие проявления. Информацию о привязанности можно получить при помощи *родительских отчетов, наблюдения в естественных условиях* или *структурированной оценки в лабораторных условиях*. В качестве примера первых двух подходов приведен *Q-набор «Привязанность»*, а в качестве примера третьего подхода — *«Незнакомая ситуация»*. В исследованиях с использованием этих процедур были выявлены важнейшие индивидуальные различия в характере привязанности детей. В дальнейших исследованиях анализировались причины этих различий и влияние их на дальнейшее развитие.

Главу завершает рассмотрение понятия *темперамента*. Под темпераментом подразумевается стиль поведения — индивидуальные особенности по таким параметрам, как уровень активности, приближение/уход и отвлекаемость. Чаще всего темперамент оценивается со слов родителей, хотя некоторые программы исследования включают метод наблюдения. Так же как и в отношении привязанности, основной вопрос, стоящий перед исследователями, касается причин индивидуальных различий, причем особое значение имеет возможность генетической обусловленности. Кроме того, интерес представляет возрастная стабильность темперамента, информацию о которой может дать лонгитюдное исследование.

Упражнения

Сегодня, одной из задач исследователей когнитивной сферы младенцев является разработка измерительных процедур, предъявляющих, в отличие от методов Пиаже, минимальные требования к моторике ребенка. Один из примеров — описанные выше исследования Байаржон, направленные на изучение постоянства объекта. Предположим, вам бы хотелось изучить понимание младенцем причинно-следственных связей с использованием процедур, аналогичных тем, которые использовала Байаржон. Какого рода эксперименты вы бы разработали?

В табл. 10.4 воспроизводится 8 из 90 утверждений *Q-опросника «Привязанность»*. Предположим, перед вами была бы поставлена задача составить вопросы для подобной методики. Какие еще поведенческие проявления и характеристики младенца, по вашему мнению, могли бы быть показательными при оценке привязанности? Составив свой список, достаньте экземпляр *Q-опросника «Привязан-*

ность» и сравните свои идеи с первоисточником (данные о котором можно найти в тексте).

Следующее упражнение подразумевает, что вы знаете по крайней мере одного, а предпочтительнее нескольких родителей с ребенком младенческого возраста, которых вы могли бы проинтервьюировать. Достаньте экземпляр одной из методик оценки темперамента, основанной на родительском отчете, либо опросник «Темперамент младенца», приведенный в таблице 10.5, либо любую другую методику, которую упоминают в своем обзоре Хуберт с соавторами или Слабач с соавторами (Hubert et al., 1982; Slabach et al., 1991). Если вы работаете с несколькими родителями, сравните две методики. Предъявите методику в соответствии со стандартной процедурой, а затем поинтересуйтесь мнением родителей о том, насколько хорошо эта методика описывает младенца.

Когнитивное развитие

Направления, на которые можно подразделить область психологии развития, столь же многочисленны и разнообразны, сколь многочисленны и разнообразны направления психологии в целом. Схема изложения, которой мы будем придерживаться в следующих двух главах, — разделение на «когнитивное» и «социальное» — наверное, наиболее обобщенное и глобальное разграничение, которое можно сделать. Эта схема не охватывает всю проблематику данной области, не всегда прозрачна и граница между когнитивным и социальным. Тем не менее такое разделение довольно типично и, как правило, имеет некоторый практический смысл.

Материал, представленный в этой главе, разбит на четыре раздела. В первой половине главы речь пойдет об исследованиях, берущих начало в двух наиболее значительных подходах к изучению детского интеллекта: о пиажетианском подходе и о тестовом, или *IQ*-подходе. Вторая половина главы посвящена двум важнейшим аспектам когнитивного развития, которые не получили должного освещения в работах представителей пиажетианской и тестовой школ: памяти детей и различного рода направлениям в развитии, которые можно объединить под рубрикой «теория сознания» (*theory of mind*).

Возрастной период, который охватывают большинство из описанных в этой главе исследований, — от 2 до 16 лет, то есть постмладенческое детство. В главе 10 мы уже рассматривали подход Пиаже к изучению когнитивного развития младенцев; в этой главе мы коснемся некоторых других форм исследований младенцев. Мы также вернемся к некоторым из представленных подходов в главе 13, говоря об исследованиях лиц пожилого возраста.

Подход Пиаже

Исследования Пиаже

В главе 10 мы отметили, что работы Пиаже занимают доминирующее положение в области исследования когнитивной сферы в младенчестве. Не меньшее влияние имеет этот подход и в области изучения постмладенческого детства. Объем литературы, посвященной этим вопросам, намного превышает масштабы исследований

Вставка 11.1

Примеры ответов на задания Пиаже о сохранении числа**ХОК (4;3):**

«Смотри. Это бутылки из кафе. Ты — гарсон и должен взять стаканы из буфета. На каждую бутылку должен быть один стакан». (Ставит точно один стакан перед каждой бутылкой и игнорирует другие стаканы). «Поровну? — Да. — (Тогда сдвигают бутылки в грудку). Бутылок и стаканов поровну? — Нет. — Где больше? — Стаканов больше». (Снова ставят стаканы точно друг перед другом, а затем сдвигают в кучу стаканы). «Одинаково стаканов и бутылок? — Нет. — Где больше? — Больше бутылок. — Почему бутылок больше? — Потому что больше».

БОК (4;7):

«Положи сюда столько же конфет, сколько их здесь. Эти (6) — для Роже. Тебе нужно взять ровно столько же. (Размещает в плотном ряду десяток конфет, но без уравнивания с десятком из ряда-модели). — Одинаково? — Нет еще (добавляет). — А теперь? — Да. — Почему? — Потому что здесь вот так (показывает длины рядов). — (Разуплотняют шесть элементов модели). У кого больше? — У Роже. — Почему? — Потому что у него доходит вот до сих пор. — Что нужно сделать, чтобы было поровну? — Добавить (прибавляет 1). — (Уплотняют эти 7 элементов и разуплотняют его ряд). — Теперь у меня больше».

Источник: *The Child's Conception of Number* (pp. 44, 75), J. Piaget & A. Szeminska, 1952, New York: Humanities.

младенчества — около 25 книг Пиаже и его коллег, а также буквально тысячи исследовательских работ на эту тему, проведенных другими специалистами. Цель этого раздела — осветить некоторые из ключевых вопросов и важнейших примеров подобных исследований. Более подробно эту тему рассматривают Флэйвелл (Flavell, 1963), Гинсбург и Оппер (Ginsburg & Oppen, 1988), Войат (Voyat, 1982) и Миллер (Miller, 1982).

Начну с примера. Во вставке 11.1 приводятся два протокола из работы Пиаже и Шеминска «Генезис числа у ребенка» (*The Child Conception of Number*, Piaget & Szeminska, 1952). В центре внимания находится такое понятие как сохранение: понимание того, что количественные характеристики объекта (набора объектов) не изменяются при изменении внешнего вида этого объекта (набора объектов). Конкретная форма сохранения, о которой идет речь, — сохранение числа: понимание того, что количество предметов при изменении их расстановки остается неизменным. Как явствует из протоколов, маленькие дети не сразу понимают принцип сохранения; как правило, они рассматривают количество предметов с точки зрения того, как эти предметы непосредственно предстают в их восприятии. Так, 4-летнему Бо ясно, что в более длинном из двух рядов должно быть больше конфет.

Мы сразу можем заметить то общее, что имеют проводившиеся Пиаже исследования младенчества с его подходом к изучению последующих периодов развития. Он вновь все внимание уделяет представлениям, которые являются неотъемлемой, фундаментальной частью процесса познания. Точно так же, как понятие об объекте представляет собой главное достижение младенца при овладении им сенсомоторного мира, понимание принципов сохранения является необходимым условием развития у старших детей мышления более высокого уровня. Между понятием об объекте и сохранением есть и более специфическое сходство. Оба они представляют важнейшие *инварианты*: аспекты окружающего мира, остающиеся неизменными, несмотря на изменения неких внешних аспектов. Всю свою карье-

ру исследователя Пиаже проявлял большой интерес к инвариантам, которые осваивает ребенок на разных этапах развития. Слова «осваивает» отражает еще одно сходство: и понятие об объекте, и сохранение, какими бы элементарными они ни казались, не даются с рождения; они должны сформироваться в период детства. Несомненно, одной из причин интереса к работам Пиаже всегда была способность этого ученого удивлять, показывая то, что дети, по крайней мере некоторое время, не понимают.

Помимо этого содержательного сходства сравнение понятия об объекте и сохранении обнаруживает общность в подходе к исследованию. Опять-таки Пиаже отказывается от строго стандартизованного подхода в пользу гибкого, ориентированного на открытия «зондирования» возможностей ребенка. И вновь Пиаже сообщает о полученных результатах преимущественно в форме протоколов, а не значений группового среднего и статистических критериев.

Между исследованиями младенчества и исследованиями периода детства есть и определенные различия. Одно из наиболее очевидных различий касается объема выборки. Как мы отметили, выборка младенцев была ограничена тремя собственными детьми Пиаже. Выборка же детей старшего возраста значительно больше и репрезентативнее. Сказать что-то более определенное довольно трудно, поскольку Пиаже редко предоставлял точную информацию об объеме и составе выборки (хотя в одной из его книг — «Развитие логического мышления в раннем детстве» (*The Early Growth of Logic in the Child*, J. Piaget & B. Inhelder, 1964) — указывается общий объем выборки равный 2159!). Умолчание о характеристиках выборки — лишь один из недостатков, которым страдают многие исследовательские отчеты. Тем не менее можно с уверенностью сказать, что выборки детей старшего возраста, которые использовал Пиаже, были значительно больше выборки младенцев.

Различия в выборках обуславливают и некоторые другие различия. Проводившиеся Пиаже исследования младенчества были лонгитюдными. За исключением нескольких работ, посвященных долговременной памяти (Piaget & Inhelder, 1973), исследования со старшими детьми проводились методом поперечных срезов. Аналогично, исследования младенчества были внутрисубъектными, в том смысле, что интересовавшие Пиаже феномены и закономерности развития изучались у всех троих малышей. Работы, посвященные периоду детства, по-видимому, практически все были межсубъектными. «По-видимому» — потому что Пиаже не всегда уточнял, основаны его выводы на внутрисубъектном или на межсубъектном сравнении.

Последнее различие касается обстановки проведения исследования. Исследования младенчества проводились в домашних условиях и были направлены в основном на изучение естественного поведения в естественных ситуациях. Суть исследований старших детей заключается преимущественно в предъявлении заданий в некоем лабораторном контексте. Как явствует из протоколов книги «Число», процедура все же может напоминать игру, а не тестирование, а взаимодействие взрослого с ребенком может быть похожим скорее на непринужденный разговор, а не школьный опрос. Однако факт остается фактом — оценивается поведение в экспериментально созданных ситуациях решения задач, а не спонтанная когнитивная деятельность. В дальнейшем мы еще вернемся к этому вопросу.

Рассмотрим еще несколько примеров, иллюстрирующих пиажетианские задания на сохранение числа. Замечу прежде, что Пиаже изучал сохранение не только

Вставка 11.2

Примеры ответов в заданиях Пиаже на сохранение веса

СЬЮЗ (6;6) изучает два шарика: «Конечно же, они весят одинаково. — А если раскатать один из них, будет ли он весить столько же? — Нужно проверить». Один из шариков раскатывают: «Нет, шарик довольно тяжелый, но тот весит немного больше, вы его раскатали, и поэтому он должен весить больше. — А можно его обратно превратить в шарик? — Да. — Он станет больше или меньше? — Я не знаю; а-а, он будет таким же, потому что он уже был шариком. — В каждом из них одинаковое количество глины? — Да. — А весят они одинаково? — Нет».

ФОГ (9;9) «Они весят одинаково. Это одинаковые шарики, вы просто вытянули один из них. — Изменился ли вес, когда я раскатал шарик? — Сначала он был круглый, а теперь продолговатый, но это та же глина; вы ничего не отщипнули. — А могу я опять скатать это в шарик, который будет весть столько же, сколько и раньше? — Конечно, можете, если не добавите глины».

Источник: *The Child's Construction of Quantities* (pp. 24, 43), J. Piaget & B. Inhelder, 1974, New York: Basic Books.

числа, но и многих других количественных параметров. В действительности, тесты сохранения содержатся в большинстве его книг, посвященных когнитивному развитию в раннем и среднем детстве. Есть исследования сохранения массы, веса и объема; длины, площади и расстояния; времени, скорости и движения. Все они воплощают в себе единый подход: предъявляются два стимула, равные по какому-либо количественному параметру; затем один из стимулов видоизменяется таким образом, что стимулы уже *не выглядят* равными, а ребенка спрашивают, равны теперь или не равны количественные параметры. Все эти исследования, кроме того, демонстрируют переход от отсутствия сохранения, обусловленного опорой на перцептивные признаки, к сохранению, основанному на логическом рассуждении. Протоколы, представленные во вставке 11.2, иллюстрируют этот переход в отношении сохранения веса.

Сохранение — это лишь одно из десятков логических и физических понятий, которые со своими сотрудниками изучал Пиаже. Ограничимся кратким описанием двух из них. Первый — **сложение классов**: принцип, согласно которому подкласс не может быть больше, чем класс более высокого порядка. Сложение классов — это, к примеру, понимание того, что маков не может быть больше, чем цветов, а уток — больше, чем птиц. Обе эти задачи, в действительности, были включены в проводившиеся Пиаже исследования классификации (Inhelder & Piaget, 1964; Piaget & Szeminska, 1952). Во вставке 11.3 приведен третий пример. Стимульным материалом являются деревянные бусины, большая часть которых — коричневые, а две — белые. Ответы 6-летней Бис свидетельствуют о том, что сложение классов, как и сохранение, — это еще одно базовое понятие, которое не присутствует с рождения, а должно развиваться.

Помимо классификации в своих исследованиях Пиаже уделял внимание и пониманию ребенком отношений между величинами. Особый интерес представляет понятие **транзитивности**. Транзитивность отражается в рассуждениях следующего рода: если, по какому-либо количественному параметру, A равно B , а B равно C , то A должно быть равно C . Или, если A больше B , а B больше C , то A должно быть больше C . Способность рассуждать подобным образом чаще всего изучалась в от-

Вставка 11.3

Пример ответов в задании Пиаже на сложение классов**БИС (6;8):**

«Больше деревянных бусинок или коричневых? — Коричневых больше, потому что есть только две белые. — А белые из дерева? — Да. — А коричневые? — Тоже. — В таком случае больше коричневых или больше деревянных бусинок? — Коричневых больше. — А какого цвета были бы бусы из деревянных бусинок? — Коричневые и белые (Как видно, Бис очень хорошо понимает исходные условия задачи!) — А бусы из коричневых бусинок? — Коричневые. — В таком случае какие бусы были бы самые длинные: бусы из деревянных бусинок или бусы из коричневых бусинок? — Из коричневых. — Нарисуй мне бусы. (Бис рисует ряд черных кружков для бус из коричневых бусинок и ряд черных кружков плюс два белых кружка для деревянных бус.) — Очень хорошо. В таком случае, какие бусы будут самыми длинными, из коричневых бусинок или из деревянных бусинок? — Из коричневых бусинок». Хорошо видно, что Бис совсем не удается разрешить эту проблему путем включения класса коричневых бусинок в класс деревянных бусинок, несмотря на точное понимание и правильное графическое представление исходных данных проблемы!

Источник: *The Child's Conception of Number* (p. 164), J. Piaget & A. Szeminska, 1952, New York: Humanities.

ношении длины и веса. Но какой бы количественный параметр ни использовался, подход всегда один: демонстрация отношений $A-B$ и $B-C$, за которым следует вопрос об отношениях между A и C . Заметьте, что количественные отношения между A и C нельзя оценить по внешнему виду; поэтому, чтобы дать правильный ответ, ребенок должен использовать результаты двух первых сравнений. Согласно исследованиям Пиаже, способность к такому логическому рассуждению развивается у детей не раньше, чем к 8–9 годам.

Задания, которые описывались до этого момента, направлены на оценку мышления в среднем детстве или, как это обозначил в своей теории Пиаже, на этапе конкретных операций. В книге под названием «Развитие логического мышления с детства и до подросткового возраста» Инельдер и Пиаже рассматривают более высокий уровень мышления, которого, по результатам их исследований, дети достигают лишь в подростковом возрасте (*The Growth of Logical Thinking from Childhood to Adolescence*, Inhelder & Piaget, 1958). Согласно теории Пиаже, этот заключительный этап называется периодом *формальных операций*. Суть формально-операционального мышления заключается в способности к гипотетически-дедуктивному рассуждению, способности абстрагироваться от непосредственной данности, целенаправленно и логически мысля категориями возможного. Аналогом такого рассуждения является научный подход к решению задач; задания Инельдера и Пиаже для изучения формальных операций фактически представляют собой задачи по физике. Одна из них, задание с маятником, использовалась в исследованиях, протоколы которых можно найти во вставке 11.4. Задача ребенка — определить, какой фактор или какие факторы влияют на частоту колебаний простого маятника. Для решения необходимо учесть все потенциально значимые переменные (вес, длину веревки, силу толчка и т. д.); целенаправленно проверять каждую из них, поддерживая при этом константность других переменных; и, наконец, сделать логические выводы на основании результатов всех проверок. Во вставке 11.4 представлены примеры неудачной попытки младшего ребенка и удачной попытки старшего ребенка решить эту задачу.

Вставка 11.4

Примеры ответов на задачу с маятником в исследовании Пиаже и Инельдера

ПЕР (10;7) — яркий пример неспособности разделить переменные: он одновременно изменяет вес и силу толчка; затем вес, силу толчка и длину веревки; затем силу толчка, вес и угол наклона и т. д., и сначала заключает: «Она зависит от веса и толчка, точно не от веревки». — «Почему ты думаешь, что веревка тут ни при чем?» — «Потому что это та же веревка». В последних попытках он не изменял ее длину; до этого он изменял ее одновременно с силой толчка, таким образом, усложняя интерпретацию результатов эксперимента. — «Но изменяется ли скорость?» — «На это влияет много всего, иногда она такая же... Да, не очень... Она также зависит от высоты, на которую вы ее [веревку] поднимаете. Когда отпускаешь ее с низкой высоты, скорость совсем небольшая». Затем он приходит к заключению, что имеют значение все четыре фактора: «Нужно изменить вес, толчок и т. д. Когда веревка короткая, он качается быстрее», но в то же время «из-за изменения веса и усиления толчка» и «из-за высоты, можно установить ее выше или ниже». — «Как ты это докажешь?» — «Можно попробовать подтолкнуть его, опустить или поднять веревку, изменить высоту и вес». Он пытается изменять все факторы сразу.

ЗМЕ (15;1), выбрав 100-граммовый груз с длинной и средней веревкой, затем — 20-граммовый груз с длинной и короткой веревкой, и, наконец — 200-граммовый груз с длинной и короткой веревкой, заключает: «Именно длина веревки заставляет маятник качаться быстрее или медленнее; вес груза не играет никакой роли». Таким же образом она отвергает факторы высоты падения и силы толчка.

Источник: *The Growth of Logical Thinking from Childhood to Adolescence* (p. 71, 75), B. Inhelder & J. Piaget, 1958, New York: Basic Books.

Рассмотрим теперь последний пример задания Пиаже. Не все исследования соответствуют логической или физической модели, как исследования, описанные выше. Первая книга Пиаже, «Речь и мышление ребенка» (*The Language and Thought of the Child*, 1926), была посвящена важнейшему компоненту знаний ребенка о социальном мире: способности встать на точку зрения другого человека. Может ли ребенок понять, что видит, думает, чувствует или желает другой человек? В частности, может ли ребенок судить об этом, когда точка зрения другого отличается от его собственной? Такое *принятие чужой точки зрения* (которое иногда называют *принятием роли другого*) является неотъемлемой частью понимания социальных отношений и социального взаимодействия. Противоположным принятию чужой точки зрения является понятие *эгоцентризма*: неспособности отвлечься от собственной точки зрения и встать на позицию другого. Думаю, учитывая рассмотренные выше когнитивные ограничения, ни у кого не вызовет удивления мнение Пиаже о том, что маленькие дети эгоцентричны.

Как следует из вышесказанного, принятие чужой точки зрения наблюдается в разных контекстах и принимает разные формы. Сам Пиаже изучал принятие чужой точки зрения в двух основных контекстах. Первый из них имел отношение к попыткам ребенка сообщить какую-либо информацию. Сообщение информации требует некоторого отвлечения от собственной точки зрения с тем, чтобы сделать доступным его для слушателя. Пиаже судил о способности к такому коммуникативному принятию чужой точки зрения как на основе спонтанных разговоров детей, так и на основе выполнения ими экспериментальных заданий, в которых требовалось донести некую информацию до другого человека (например, пересказать

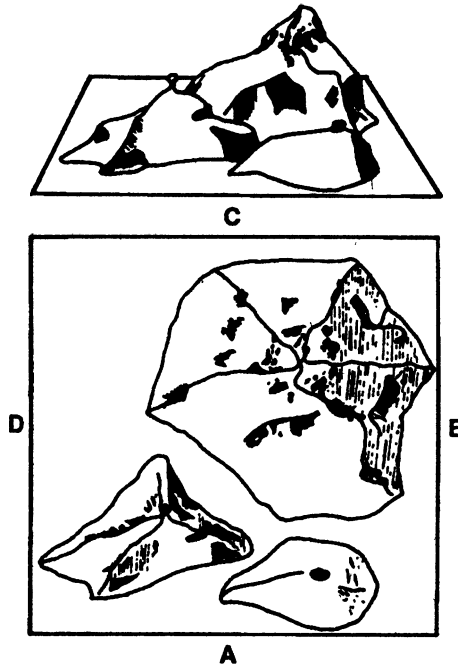


Рис. 11.1. Модель «три горы» для оценки визуального принятия чужой точки зрения. (Источник: *The Child's Conception of Space* (p. 211), J. Piaget & B. Inhelder, 1956, London: Routledge & Kegan Paul)

только что рассказанную историю). Второй контекст имел отношение к пространственному принятию чужой точки зрения, или способности представить, что видит другой человек. Главным средством изучения пространственного принятия чужой точки зрения была модель «три горы», изображенная на рис. 11.1. После того как ребенок обходил вокруг макета, его усаживали с одной из сторон модели и просили описать, что увидит кукла, если ее посадить на ту или иную сторону.

Я вернусь к некоторым из описанных выше заданий, когда речь пойдет об исследованиях, направленных на проверку теории Пиаже. Но прежде чем сделать это, стоит сказать еще несколько слов о подходе Пиаже к исследованию. Процедура интервью, которую иллюстрируют протоколы, представленные в этой главе, называют **клиническим методом** тестирования. Пиаже (Piaget, 1929) использовал термин «клинический метод» по причине сходства своего подхода с подходом профессионального клинициста, занимающегося постановкой диагноза и коррекцией эмоциональных нарушений. Для обоих методов характерна гибкость: исследователь имеет возможность отклониться от строгой процедуры для проверки реакции конкретного испытуемого разнообразными, не заданными заранее способами. В умелых руках клинический метод может стать великолепным инструментом разработки новых областей исследования. Он также может отлично выполнять диагностическую функцию, позволяя точно определять, каков в действительности взгляд ребенка на окружающий мир, что бывает затруднительно при неотступном следовании стандартной процедуре, которую нельзя подчинить индивидуальным осо-

бенностям (недостаточному словарному запасу, временному переключению внимания, своеобразию интерпретаций и т. д.) отдельного ребенка.

Самое полное описание клинического метода Пиаже дал в одной из первых своих книг, «Представления о мире у ребенка». Нижеследующий отрывок стоит того, чтобы привести его целиком:

«По нашему мнению, в детской психологии, как и в патопсихологии, для преодоления этапа блуждания в потемках, с которым неизбежно сталкивается любой новичок, необходимо по крайней мере год ежедневной практики. Так трудно, особенно педагогу, не говорить слишком много, расспрашивая ребенка! Так трудно не навязывать свое мнение! И главное, так трудно найти средний путь между излишней систематизацией по причине наличия уже сформировавшихся представлений и совершенной непоследовательностью по причине отсутствия какой-либо направляющей гипотезы! Фактически хороший экспериментатор должен соединять в себе два часто несовместимых качества; он должен знать, как наблюдать, то есть позволять ребенку свободно излагать свои мысли, не обрывая и не перебивая его; и в то же время он должен постоянно вычленять нечто определенное, у него всегда должна быть рабочая гипотеза, некая теория, верная или ошибочная, которую он пытается проверить. Чтобы оценить реальную трудность клинического метода, нужно научить ему кого-то другого. Начинаящие студенты либо внушают ребенку все, что они надеются получить, либо не внушают ничего, поскольку не стремятся ничего выявить, а это верный способ ничего не обнаружить» (Piaget, 1929, p. 7–8).

Клинический метод — одна из важнейших составляющих исследований Пиаже, однако он не имел бы большого значения, если бы не существовало действительно интересных конструкторов, для изучения которых его можно было бы применить. И здесь мы подходим к рассмотрению, наверное, самой сильной стороны подхода Пиаже: речь идет о широчайшем диапазоне форм знаний и изощренных способах оценки этих знаний. По подборке заданий и результатов исследований, включенных в эту книгу, можно получить лишь отдаленное представление об этом разнообразии. Книги Пиаже сохранили новизну и научную ценность, необходимые для изучения того, как рассуждают дети, — они содержат описание методов, послуживших образцом для методов, используемых в значительной части исследований когнитивного развития. В главе 1 я отмечал, что техничность проведения исследования должна всегда сочетаться с интересными идеями о том, что и как изучать. Вряд ли у кого-либо другого в истории психологии развития было столько же интересных идей, сколько у Пиаже.

Спорные вопросы и дальнейшие исследования

Перейдем теперь от похвал к критике. Проверке положений теории Пиаже посвящено огромное количество исследований, причем основанием для проведения многих из них послужили явные просчеты самого Пиаже. Эти более глубокие исследования будут рассмотрены под тремя рубриками: оценка, закономерности и когнитивные изменения.

Вопрос *оценки* — это та же основополагающая проблема, о которой шла речь при обсуждении работ Пиаже, посвященных младенчеству: можно ли с помощью процедур, использовавшихся Пиаже, получить точное представление о способностях ребенка? Как и в отношении исследований младенчества, наибольшее беспокой-

ство вызывает возможность недооценки истинных знаний маленького ребенка. Хотя было выделено множество факторов, которые могут привести к ошибке при вынесении диагностического решения, чаще всего поводом для критики становится чрезмерный акцент на вербальном компоненте во многих заданиях Пиаже. Это очевидно из протоколов, представленных ранее в этой главе. Рассмотрим задание на сохранение. Оно предназначено для оценки понимания ребенком логики, лежащей в основе сохранения. Однако о понимании ребенком этой логики нельзя судить по каким-то внешним проявлениям; его можно оценить только с использованием языка, как в вопросах, задаваемых ребенку, так и в ответах, которые должен дать ребенок. Представляется довольно правдоподобным, что ребенок, который якобы «не понимает принципов сохранения», в действительности просто путает такие слова как «одинаковый», «больше», «меньше» и «количество». Возможно, к примеру, что ребенок относит слово «больше» к длине ряда, а не к количеству предметов.

Для исследования этой возможности были использованы разнообразные подходы. Одни исследователи применяли вербальные претесты для оценки способности ребенка понять слова, которые используются в заданиях на сохранение (например, Miller, 1977). Другие — не ограничились претестом, а пошли дальше и использовали предварительное вербальное обучение, то есть обучение ребенка перед тестированием словам, используемым в тесте (например, Gruen, 1965). А некоторые исследователи попытались справиться с потенциальной путаницей в словах, разработав для оценки пиаже-тианских понятий «невербальные» процедуры. Такие процедуры редко бывают невербальными в буквальном смысле этого слова; однако они проще с точки зрения лексики, и в них не используются такие потенциально проблематичные слова, как «одинаковый» или «больше». Одна из методик направлена на оценку понимания ребенком относительного количества посредством процедуры, в которой ребенку предлагается выбрать между двумя наборами конфет, один из которых мог бы показаться большим (незрелый ответ), а другой был в действительности больше (зрелый ответ) (например, Miller, 1976a). Еще одна стратегия — искусственное создание ситуаций нарушения неких принципов, лежащих в основе изучаемого понятия (например, создание видимости нарушения принципов сохранения) с последующим измерением реакции ребенка; при этом предполагается, что реакция удивления свидетельствует о том, что испытуемый имеет некоторые представления об изучаемом понятии (например, Gelman, 1972). Вспомним, что подход по типу обмана ожиданий использовался при изучении понимания младенцами постоянства объекта.

Не вся полемика была связан с проблемой языка. Сомнения вызывает и общий контекст, в рамках которого обычно оцениваются представления о сохранении. Этот контекст, по сути, довольно непривычен для ребенка и включает ряд особенностей, которые могут подтолкнуть к даче неверного ответа на вопрос о сохранении. Среди этих особенностей: явный интерес к количеству, кажущаяся необоснованность производимых трансформаций с предметами (почему взрослый раздвигает конфеты?) и необходимость в короткий период времени дважды отвечать на один и тот же вопрос, что может восприниматься маленькими детьми как сигнал к изменению своего ответа. Возможно, если бы обстановка была более естественной

и знакомой, у ребенка было бы меньше шансов быть причисленным к тем, кто не понимает принципов сохранения.

Эта возможность исследовалась разными способами. Роуз и Блэнк (Rose & Blank, 1974) изучали влияние повтора одного и того же вопроса о сохранении путем простого исключения одного вопроса при работе с половиной детей. Макгарригл и Дональдсон (McGarrigle & Donaldson, 1974) заменили намеренное произведение трансформаций взрослым экспериментатором внешне случайным изменением, вносимым «непослушным» мишкой, который, по мнению исследователей, должен быть лучше знаком ребенку и не восприниматься им как некто, наделенный магической властью изменять количество. Лайт, Бакинхэм и Роббинс (Light, Buckingham, & Robbins, 1979) производили аналогичные манипуляции, однако у них трансформация носила второстепенный характер, а не воспринималась как неприятная неожиданность, то есть цель производимых трансформаций не сводилась к изучению сохранения, они вводились естественным путем в ходе игры.

До этого момента ничего не было сказано о результатах таких исследований с модифицированной процедурой оценки. Хотя интерпретации их довольно противоречивы, разумными кажутся три общих вывода (более подробно об этом см. у Chandler & Chapman, 1991; Miller, 1976b и Siegal, 1991). Первый — использование методов Пиаже действительно приводит к недооценке способностей маленького ребенка, поскольку с модифицированными тестами дети часто справляются лучше, чем со стандартными тестами Пиаже. Второй — недооценка эта не слишком велика; такие феномены, как непонимание принципов сохранения нельзя полностью объяснить вербальными затруднениями или контекстуальными искажениями. Третий — процесс развития таких когнитивных навыков, как сохранение и принятие чужой точки зрения, более длительный и сложный процесс, чем это представлялось Пиаже, с множеством элементарных уровней и предпосылок, к которым нечувствительны процедуры Пиаже. В основе этого вывода не только результаты описанных выше исследований, но и данные, полученные в исследовательских программах, направленных на изучение более простых навыков, формирующихся на более ранних этапах развития, чем те навыки, которые оценивал Пиаже. В этой связи стоит особенно отметить работы по принятию чужой точки зрения Джона Флэйвелла (Flavell, 1992) и работы по представлениям о числе Рошель Гельман (Gelman, 1991).

Второй вопрос, который мы рассмотрим, касается *закономерностей* развития. Вопрос закономерностей — это вопрос о том, как соединяются друг с другом различные когнитивные навыки. Согласно теории Пиаже, существуют два основных типа развития когнитивных навыков. Первый — инвариантная последовательность: появление двух или более навыков в одном порядке у всех детей. Теоретически инвариантность связана с возможностью того, что навыки, формирующиеся на ранних этапах, служат в качестве необходимых посредников или фундамента для освоения навыков более высокого уровня. Второй тип — синхрония: появление двух или более навыков одновременно, на одном этапе развития. Здесь с теоретической точки зрения интересна возможность общего базиса разных навыков.

С позиции методологии изучение последовательностей и синхронии должно отвечать двум требованиям. Во-первых — внутрисубъектное тестирование. На это

указывалось в главе 3 при обсуждении меж- и внутрисубъектных планов в целом. Выявление взаимосвязей и взаимоотношений в процессе развития возможно только в том случае, когда имеются данные о выполнении всех сравниваемых заданий всеми испытуемыми. При внутрисубъектном тестировании о последовательности или синхронии можно судить по соотношению успешных попыток выполнения и неудач. В табл. 11.1 представлены условия истинности гипотез об инвариантной последовательности и синхронном появлении навыков. Логика, используемая в случае с двумя навыками, применима и при наличии трех навыков и более. Например, если предполагается, что *A*, *B* и *C* появляются последовательно, возможен союз четырех результатов: неспособность выполнить ни одно из заданий, выполнение только *A*, выполнение только *A* и *B* или выполнение всех трех заданий.

Таблица 11.1

Варианты соотношения успешных и неудачных попыток выполнения заданий и их связь с последовательностью и синхронией

Предполагаемая связь	Конгруэнтные соотношения	Неконгруэнтные соотношения
Последовательность <i>A</i> и <i>B</i>	$A+ B+, A+ B-, A- B-$	$A- B+$
Синхрония <i>A</i> и <i>B</i>	$A+ B+, A- B-$	$A+ B-, A- B+$

Примечание: + означает успешное выполнение; — означает неудачную попытку выполнения.

Второе требование при исследовании закономерностей — уравнивание чувствительности заданий, используемых для сравнения разных навыков. Определение относительной трудности навыков *A* и *B* возможно только в том случае, когда тесты, используемые для измерения *A* и *B*, равно чувствительны, то есть обладают равной способностью выявлять истинные представления ребенка. Если тесты не удовлетворяют требованию равной чувствительности, существует вероятность прийти к ошибочным выводам о порядке этапов развития. Допустим, выбранный нами тест для измерения когнитивного навыка *A* носит упрощенный характер и поэтому выявляет самые зачатки формирующегося навыка. Допустим, что наш тест для навыка *B* значительно менее чувствителен, к примеру, в отличие от *A*, сложен в вербальном отношении и предъявляет высокие требования к памяти. Мы вполне можем обнаружить, что *A* удастся выполнить раньше, чем *B*, и вероятно даже, что это принимает форму некой инвариантной последовательности. Однако истинные отношения между *A* и *B* могут совершенно не соответствовать тому, на что указывают результаты использования наших методик.

В том виде, в котором мы ее здесь кратко описываем, проблема чувствительности кажется чем-то, что можно довольно легко разрешить, просто не включая задания, для выполнения которых требуются неодинаковые дополнительные усилия. Однако, как и многие методологические предписания, это легче декларировать, чем осуществить на практике. Довольно трудно бывает определить, что является «дополнительным»; в конечном счете, каждый навык заключен в некий контекст, с определенными стимулами, инструкциями, требованиями к памяти и т. д. Если, к

примеру, мы оцениваем понимание принципов сохранения, нам придется решить помимо всего прочего, что использовать в качестве стимульного материала, как установить первоначальное равенство, какие произвести трансформации, как сформулировать вопрос о сохранении и просить ли ребенка объяснить свой ответ. Заблуждение думать, что такое понятие как сохранение можно измерить в его «чистой» форме, при отсутствии подобного контекста. Заблуждение также думать, что понятия, представляющие интерес для психологии развития, — это самостоятельные, гомогенные единицы — самостоятельный навык сохранения или транзитивности или принятия чужой точки зрения, ожидающие, пока их выявит «подходящий» тест. Напротив, как было подчеркнуто при рассмотрении вопроса оценки, любой конструктор содержит ряд уровней и составляющих его навыков, уровней и навыков, не всегда одинаково проявляющих себя в разных контекстах. Все эти соображения в значительной мере усложняют задачу выявления взаимосвязей в ходе развития.

Насколько удачно разрешаются описанные выше трудности в исследованиях Пиаже? Краткий, но откровенный ответ — не слишком удачно. Несомненно, Пиаже осознавал проблему чувствительности тестов; в действительности, в его книге о числе есть красноречивое обсуждение этого вопроса (Piaget & Szeminska, 1952, р. 149). Однако в его собственных исследованиях делаются лишь ограниченные и зачастую неубедительные попытки уравнивать чувствительность заданий при сравнении разных когнитивных навыков. Возможно, даже большее значение имеет вторая из упомянутых выше проблем — отсутствие внутрисубъектных сравнений. За редкими исключениями выводы Пиаже о закономерностях постмладенческого развития основаны на показателях среднего возраста освоения когнитивных навыков в разных выборках детей, выполнявших разные задания. Заключение о наличии последовательности $A \rightarrow B$, к примеру, могло быть сделано на основании того, что одна группа детей осваивает задание A в возрасте около 6 лет, а другая группа детей осваивает задание B примерно в 8-летнем возрасте; то, что обе группы справляются с заданием в возрасте около 8 лет, могло рассматриваться как свидетельство синхронии. Ясно, что, используя такой подход, невозможно получить достаточно убедительных доказательств в отношении закономерностей развития.

После Пиаже исследования закономерностей развития стали проводиться по внутрисубъектному плану, с использованием множества задач и сложных статистических методов анализа взаимосвязей (например, Kofsky, 1966; Toussaint, 1974). В этих исследованиях также предпринимались более серьезные попытки уравнивания чувствительности заданий. Однако из сказанного выше понятно, что проблема чувствительности заданий не имеет легкого или единственно верного пути решения. На практике существует масса подходов, что ведет к противоречивым мнениям о том, каковы именно закономерности развития. Обобщенный вывод из всех этих исследований состоит в том, что многие последовательности, выделенные Пиаже, действительно имеют место; однако параллельность в развитии обнаружить значительно сложнее (более подробно об этом см.: Flavell, Miller, & Miller, 1993).

Последний вопрос, который следует рассмотреть, — это вопрос *когнитивных изменений*. Как ребенок осваивает все те навыки, которые, как явствует из работ

Пиаже, он должен освоить? Как, к примеру, осуществляется переход к пониманию принципов сохранения, переход от изначального эгоцентризма к умению встать на позицию другого человека? Как, в целом, ребенок переходит от одной стадии когнитивного развития к другой? И, конечно, с точки зрения методологии особый интерес представляет следующий вопрос: как изучать когнитивные изменения — не только овладение навыками, о которых говорил Пиаже, но и изменения, касающиеся любого аспекта когнитивного развития ребенка?

Пиаже часто критикуют за то, что он рассуждает об этапах, через которые проходит ребенок, больше, чем о механизмах этого перехода. Эта критика довольно справедлива. Проблема не в том, что Пиаже не уделял должного внимания когнитивным изменениям; эта тема достаточно глубоко проработана в его теории. Проблема в том, что между теорией и данными, полученными в соответствующих исследованиях, нередко оказывается серьезный пробел. Фактически даже не всегда ясно, какие данные могли бы иметь значение или с помощью каких методик их можно было бы собрать.

Здесь уместно вспомнить об одном из различий между исследованиями младенчества и исследованиями более позднего детства. Исследования младенчества проводились в естественных условиях и по лонгитюдному плану; такое сочетание позволяло Пиаже наблюдать ребенка в процессе преодоления все новых и новых трудностей и перехода, зачастую очень медленного и постепенного, на новые уровни понимания. Вполне вероятно, что многие идеи о когнитивном развитии сформировались у Пиаже в то время, когда он наблюдал за тем, что происходит с его собственными маленькими детьми. Исследования с детьми старшего возраста не были ни натуралистическими, ни лонгитюдными; лишь несколько из них проводились методом непосредственного наблюдения за детьми, осваивающими нечто новое. Практически все исследования, не связанные с младенчеством, носят *диагностический* характер, то есть направлены на выявление когнитивных навыков, которые у ребенка уже сформировались.

Поэтому я обращаюсь к работам других специалистов, занимавшихся вопросом когнитивных изменений. На сегодня самой популярной парадигмой изучения когнитивных изменений является *обучающее исследование*. Как следует из названия, цель обучающего исследования — научить ребенка тому понятию, которое он еще не освоил. Чаще всего в таких исследованиях используется понятие сохранения, поэтому в качестве примера мы его и рассмотрим. Первая фаза обучающего исследования состоит в претесте сохранения для определения, кто из детей уже понимает принципы сохранения, а кто — не понимает и поэтому является кандидатом на обучение. Затем выявленные таким образом дети переходят на вторую, обучающую фазу исследования. На этой фазе детей подвергают различным экспериментальным воздействиям, цель которых — сформировать представления о сохранении. В чем конкретно будет состоять это воздействие, в значительной мере зависит от теоретических предпочтений исследователя; однако общее для всех исследований — стремление сделать так, чтоб экспериментальные воздействия имели некоторое отношение к реальным жизненным ситуациям или процессам, посредством которых дети осваивают принципы сохранения. Заключительная фаза — посттест сохранения для определения эффективности обучения. Если обучение

оказывается успешным, по его эффективности мы в некоторой степени можем судить о реальном процессе освоения принципов сохранения. Следует подчеркнуть, что возможности этой стратегии выходят далеко за пределы ее применения для изучения феноменов Пиаже: проверяется гипотеза о том, что *A* приводит к *B*, для этого экспериментально создают *A* и смотрят, появилось ли *B*.

Существует огромное количество обучающих исследований. Не все они были успешными; в действительности, первые такие исследования, проведенные в конце 1950-х — начале 1960-х, оказались в основном неудачными (оценку первых работ в этой области можно найти в Flavell, 1963). Однако в последнее время общая тенденция изменилась, и сегодня уже нет сомнений, что принципам сохранения *можно* успешно обучать — не всех детей, конечно, но, по крайней мере, некоторых из них (Field, 1987). Более того, абсолютно ясно, что к успеху могут привести процедуры, которые, на первый взгляд, не имеют ничего общего. В табл. 11.2 обобщены некоторые из методов обучения, использование которых в лабораторных условиях, как было доказано, приводит к освоению принципов сохранения.

Какие факторы следует принимать во внимание при оценке обучающего исследования? Необходимо поставить по крайней мере три вопроса.

Первый из них касается точности диагностического вывода. Действительно ли дети не понимали принципов сохранения до начала исследования? И, что важнее всего, действительно ли они стали их понимать по окончании обучения? Добиться от ребенка ответа «одинаковые» на вопросы о сохранении не представляет особого труда; весь вопрос в том, привело ли обучение к истинному пониманию соответствующего понятия, к пониманию, сопоставимому с тем, к которому приходят в процессе естественного развития, без специального обучения. Приверженцы концепции Пиаже, прежде чем признать истинность достигнутых улучшений, как правило, проверяют их на соответствие жестким критериям. Эти критерии включают в себя наличие не только верного ответа, но и его логическое обоснование; перенос навыков решения задач на другие задания подобного типа и устойчивость тенденции давать правильные ответы (к примеру, в отсроченном посттесте).

Второй вопрос касается оснований для успеха обучения. Все исследователи, конечно, имеют некоторые теоретические или эмпирические обоснования созданной ими программы обучения, определенный набор предполагаемых ситуаций или процессов, задействованных в обучении. В табл. 11.2 приведены теоретические конструкторы (обратимость, внимание и т. д.), лежащие в основе указанных программ обучения. Проблема состоит в том, что между конструктором (например, обратимостью) и экспериментальными манипуляциями (например, восстановлением первоначального численного соответствия между куклами и игрушечными кроватками после проведения трансформаций с одним или другим набором) нет прямой однозначной связи. Всегда имеет место некоторый «логический скачок» от манипуляций к объяснению, который означает, что успех любой процедуры обучения всегда имеет альтернативные объяснения. Тщательное планирование исследования может сократить количество таких альтернатив, но оно не может их полностью исключить. То есть мы часто знаем, что обучение достигло своей цели, но не знаем, почему это произошло. Это, как вам уже известно, вопрос *конструктивной валидности*. Правильно ли мы интерпретируем полученные результаты?

Таблица 11.2

Примеры пиажетианских обучающих исследований

Исследование	Формируемые понятия	Тип обучения	Процедура
Beilin (1965)	Сохранение числа, сохранение длины	Вербальное инструктирование	Вербальное объяснение правил сохранения, следующее за каждым неправильным ответом — например: «Теперь я их передвигаю. Смотри, они стоят по-другому, но их столько же, сколько было до этого. Они просто иначе выглядят. Смотри, я могу поставить их снова так же, как они стояли до этого; видишь, их столько же, сколько и прежде, поскольку я ничего не добавил и ничего не убрал. Я просто передвинул их»
Wallach, Wall, & Anderson (1967)	Сохранение числа	Обратимость	Демонстрирование того, что произведенные трансформации обратимы и количество предметов при этом остается равным. В качестве стимульного материала использовались куклы и игрушечные кроватки; в ходе обучения показывалось, что для каждой куклы всегда находится кроватка, вне зависимости от того, как расставлены предметы
Gelman (1969)	Сохранение числа, сохранение длины	Внимание к количеству	Процедура формирования установки на обучение, направленная на привлечение внимания к главным количественным атрибутам (числу или длине) и отвлечение внимания от внешних атрибутов. Обучение состоит из серии попыток, когда ребенок должен выбрать из трех стимулов два, которые были бы «одинаковыми»; подкреплялся выбор, сделанный на основе количественных атрибутов (например, числа), а не перцептивных характеристик (например, широкой или плотной расстановки)
Bucher & Schneider (1973)	Сохранение числа, сохранение массы, сохранение объема	Оперантное обусловливание	Подкрепление (похвалой или жетоном) правильных ответов при решении последовательности задач возрастающей трудности. Задачи — разнообразные по форме и сложности
Zimmerman & Rosenthal (1974)	Сохранение числа, сохранение длины, сохранение площади	Научение через наблюдение (моделирование)	Демонстрирование взрослого, правильно решающего задачи на сохранение, сходные с теми, которые получают дети. Моделирование включает как правильные ответы, так и адекватное объяснение (например: «Потому что сначала их было равное количество»)

Последняя проблема — это проблема *внешней валидности*. В лучшем случае обучающее исследование демонстрирует то, что некоторое воздействие может привести к определенному когнитивному изменению; но оно не доказывает того, что в реальной жизни такое воздействие является необходимым или обычным условием когнитивного развития. Проблема в том, что все манипуляции, направленные на обучение, связаны с опытом или ситуациями, которые отличаются от естественного опыта ребенка. Часто это отличие весьма резкое. Разумеется, такому намеренному созданию неестественной обстановки есть свое объяснение, общее для всех экспериментально-лабораторных исследований: возможность строгого контроля и точного измерения переменных, а следовательно, и высокой степени внутренней валидности при выявлении причинно-следственных связей. Исследователи, которые экспериментируют с детьми в рамках учебного процесса, не претендуют на создание зеркального отражения реальной жизни; они пытаются насытить период обучения соответствующим опытом, а также выявить и осветить процессы, играющие решающую роль. Тем не менее разница между «могут» и «делают», о которой говорилось в главе 6, справедлива и для этого случая. Обучающие исследования позволили обнаружить разного рода опыт, который *может* спровоцировать когнитивные изменения. Однако остается неясным, какого именно рода опыт *в действительности* имеет значение для нормального когнитивного развития без специального обучения.

Подход, основанный на тестировании интеллекта

Сущность IQ-тестов

Тестовый, или IQ-подход, к интеллекту во многом отличается от подхода Пиаже. Поскольку до сих пор основное внимание было уделено Пиаже, начну с рассмотрения некоторых из этих отличий.

Пиаже всегда интересовался исключительно закономерности развития, то есть общее в развитии всех детей. Например, такое понятие, как сохранение в конечном счете осваивают все нормальные дети. Если брать пример более высокого порядка, практически все дети достигают стадии конкретных операций. Индивидуальные различия заключаются в разных темпах развития; у Пиаже эти различия никогда не вызывали интереса. Смысл использования IQ-тестов, напротив, сводится к выявлению индивидуальных различий между детьми. Такие тесты, кроме того, измеряют не просто различия, а их *количественные* характеристики — мы говорим, что у одного ребенка интеллект «выше» или «ниже» среднего. Оценочный компонент является неотъемлемой частью любого теста интеллекта. Тот факт, что эти тесты заставляют нас выносить оценочные суждения о детях, — одна из причин, по которой их использование всегда вызывало многочисленные споры.

С самого начала в своих работах Пиаже руководствовался теорией, при этом стремясь ответить на основополагающие эпистемологические вопросы, касающиеся представлений ребенка о таких понятиях, как число, пространство, время и причинность. Каково бы ни было практическое применение его выводов (например, влияние на школьные программы), оно началось позже; сам Пиаже никогда не был инициатором этого применения. IQ-тесты, напротив, с момента своего создания носили прагматический характер. Первый удачный тест интеллекта был

разработан в Париже в 1905 году Бине и Симоном с исключительно прагматической целью прогноза школьной успеваемости детей. В действительности, именно возможность предсказать успеваемость явилась четким критерием при отборе заданий для теста. Начиная с того времени *IQ*-тесты имеют ту же практическую основу и то же практическое назначение. Да, у Бине была некая концепция интеллекта, которая легла в основу его теста, и в конечном счете были разработаны теории того, что же в действительности измеряют подобные тесты. Однако большинство теорий появилось вслед за практическими методиками. То есть последовательность была обратной той, которая имела место в случае с Пиаже: сначала практическое применение и только потом теория.

Последнее отличие связано с акцентом *IQ*-тестов на количественной оценке. *IQ*-тесты призваны отвечать на вопрос «сколько», а не «как». Тест дает некое число, по которому судят о том, каков уровень интеллекта ребенка. В некоторых тестах выводятся несколько показателей, соответствующих разным типам интеллекта; однако по сути своей подход остается количественным. Все, что должно интересовать при подсчете результатов *IQ*-теста, — это количество правильных ответов, которые дал ребенок; а все, что делают с правильными ответами, — это суммируют их для получения итоговой оценки. Таким образом, здесь важнее всего продукт когнитивной деятельности, а не лежащие в его основе процессы. Для приверженцев подхода Пиаже, напротив, наибольший интерес представляет процесс, а не продукт. В пиажетианских исследованиях осуществляется попытка выйти за пределы правильного или неправильного ответа ребенка для выявления качественных характеристик его системы мышления и качественных изменений, которые претерпевает эта система в ходе развития. Таким образом, здесь важнее всего не «сколько», а «как».

До этого момента все, что мы говорили об *IQ*-тестах, имело негативную окраску. Они дают количественную, а не качественную оценку, в них все внимание уделяется продукту, а не лежащим в его основе процессам, они преследуют прагматические, а не теоретические цели. Это не единственное, за что критикуют *IQ*-тесты, я еще не упомянул о том, что чаще всего критики указывают на необъективность, которая снижает шансы на успех у определенных групп населения. Учитывая все эти проблемы, естественно задаться вопросом: зачем вообще принимать всерьез *IQ*-тесты? На каком основании был сделан вывод о том, что они измеряют интеллект? Что, иными словами, свидетельствует об их *валидности*?

Вопрос валидности теста в общем виде рассматривался в главе 4. Вспомним, что валидность теста обычно доказывается наличием корреляции между результатами этого теста и результатами других методик, с которыми данный тест должен был бы быть связан, либо практическим прогнозом, либо теоретической убедительностью. Валидизация *IQ*-тестов основана на силе этой корреляционной связи. Первый подобный тест, тест Бине и Симона, был признан удачным, поскольку позволял разделить детей на тех, у кого успеваемость в школе будет, скорее всего, высокой, и тех, у кого она, скорее всего, будет низкой. Со времени создания Бине и Симоном их теста корреляция с уровнем школьной успеваемости или результатами тестов академических достижений является главным индексом валидности тестов интеллекта для детей. Обычно значение коэффициента корреляции колеблется в районе 0,5–0,6, что указывает на умеренно сильную, но далеко не абсолютную связь. Однако корреляционные возможности *IQ* не ограничиваются контекстом

академических успехов. *IQ* также коррелирует с профессиональным статусом в период взрослости и успешностью выполнения множества когнитивных тестов (Jensen, 1981). Именно эта (хотя и несовершенная) способность предсказывать уровень успехов во многих областях, в которых, очевидно, необходим достаточно высокий интеллект, является доказательством валидности *IQ*-тестов как методик измерения интеллекта¹.

Вспомним теперь то, о чем говорилось в главе 4, — стандартизованные тесты должны отвечать требованию не только валидности, но и надежности. Наибольшее значение здесь имеет ретестовая надежность: если мы дважды за непродолжительный период времени предложим один и тот же тест, получим ли мы близкие результаты? Ответ на этот вопрос в отношении *IQ*-тестов следует из того, что было сказано о валидности. Надежность — необходимое условие валидности, и поэтому доказательство того, что тест валиден, подразумевает, что он также и надежен. Фактически, ретестовая надежность основных *IQ*-тестов для детей довольно высока; коэффициент корреляции между результатами первого и второго тестирования обычно составляет около 0,9.

Примеры тестов

Многие тесты, как предполагается, измеряют интеллект; они отличаются по нескольким параметрам. Некоторые тесты дают один итоговый показатель как индекс глобального, или общего интеллекта. Наверное, наиболее известный тест общего интеллекта — это тест Стэнфорд—Бине (Thorndike, Hagen, & Sattler, 1986), «прямой исторический потомок» теста, созданного Бине и Симоном. Другие тесты измеряют более специфические способности. При работе с детьми, например, часто используется словарный тест в картинках Пибоди (*Peabody Picture Vocabulary Test*; Dunn & Dunn, 1981), позволяющий оценить устный рецептивный словарный запас. С помощью разнообразных тестов Векслера (Wechsler, 1981, 1989, 1991) дается оценка вербального и невербального интеллекта, а также итоговая оценка *IQ*.

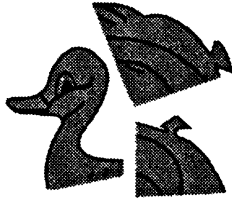
Тесты отличаются также то, на какие возрастные группы они рассчитаны. Три шкалы Векслера созданы для трех разных возрастных групп: Векслеровская шкала интеллекта для дошкольников и младших школьников (*Wechsler Preschool and Primary Scale of Intelligence*), или *WPPSI* используется при работе с детьми 4–6 лет; Векслеровская шкала интеллекта для детей (*Wechsler Intelligence Scale for Children*) или *WISC* рассчитана на возраст от 6 до 16 лет, а Векслеровская шкала интеллекта взрослых (*Wechsler Adult Intelligence Scale*), или *WAIS* предназначена для взрослых. Тест Стэнфорд—Бине охватывает более широкий возрастной диапазон; хотя обычно этот тест используют для оценки интеллекта детей, он пригоден для возрастной группы от 2 лет и старше, весь период взрослости. Наконец, хотя младенцы — это единственная группа, не охваченная тестами Векслера и Стэнфорд—Бине, есть методики, разработанные специально для оценки развития в младенчестве, — например, шкалы развития младенцев Бейли (*Bayley Scales of Infant Development*; Bayley, 1993).

¹ В действительности чаще всего в качестве индекса валидности используются коэффициенты корреляции с другими тестами интеллекта. Новый тест, к примеру, будут воспринимать всерьез только в том случае, если его результаты в достаточной мере коррелируют с результатами таких проверенных тестов, как Стэнфорд—Бине или *WISC*. Аналогично, новая версия любого проверенного теста обосновывается преимущественно через корреляцию с предыдущей версией этого теста.

Последнее различие касается метода предъявления. Одни тесты предполагают индивидуальное предъявление — то есть на одного тестирующего приходится один тестируемый. К таким методикам относятся все упомянутые выше тесты интеллекта детей и взрослых: Стэнфорд—Бине, Векслер и Пибоди. Естественно, сюда входят и тесты интеллекта младенцев. Другая категория тестов предназначена для группового предъявления, что означает возможность в одно и то же время тестировать множество людей. Групповые тесты интеллекта приобрели особое значение в период Первой мировой войны, когда они использовались для оценки интеллекта новобранцев. С тех пор групповые тесты стали довольно часто использоваться и в школах. Эти тесты имеют очевидное преимущество в связи с легкостью предъявления и эффективностью; за время, которое бы понадобилось на оценку интеллекта одного ребенка с помощью методики Стэнфорд—Бине, можно протестировать 30–40 детей. Групповые тесты имеют также и очевидные недостатки: их можно использовать только с испытуемыми, достаточно зрелыми, чтобы отвечать на вопросы в письменном виде; и, кроме того, испытуемые не всегда выполняют групповые тесты на оптимальном для себя уровне.

Таблица 11.3

Типы заданий, включенных в Векслеровскую шкалу интеллекта для детей, третье издание

Субтест	Задания
Осведомленность	Сколько крыльев у птицы? Сколько в десяти центах пятицентовых монет? Что такое перец?
Арифметический	У Сэма три конфеты; Джо дал ему еще четыре. Сколько всего конфет стало у Сэма? Если два яблока стоят \$.15, сколько будет стоить дюжина яблок?
Словарный	Что такое _____? или Что означает _____? Молоток Предохранять Эпидемия
Субтест	Невербальная шкала
Складывание картинок	Составьте из нескольких частей картинку с изображением знакомого предмета 

Пока практически ничего не было сказано о содержании *IQ*-тестов. В табл. 11.3 приведены примеры заданий, включенных в один из основных тестов интеллекта для детей: в Векслеровскую шкалу интеллекта для детей, или *WISC*. Другой тест интеллекта детей, также занимающий одно из ведущих мест, — тест Стэнфорд—Бине — имеет много общего с *WISC*. Оба теста состоят из ряда субтестов, предназначенных для оценки разных способностей — 11 субтестов в *WISC* и 15 субтестов в Стэнфорд—Бине. В обоих тестах субтест можно сгруппировать в глобальные шкалы: «Вербальную» и «Невербальную» шкалы в *WISC*; и в шкалы «Вербальное рассуждение», «Количественное рассуждение», «Абстрактное/наглядное рассуждение» и «Кратковременная память» в Стэнфорд—Бине. В этих методиках уделяется внимание сходным видам когнитивных способностей. Даже задания, которые не носят явно вербального характера (например, невербальные тесты в *WISC*), обычно предполагают понимание ребенком словесных инструкций. Важна также память, как память на осмысленный материал, так и механическая память (оба теста, к примеру, включают задания на запоминание последовательности цифр). Ряд субтестов направлен на оценку арифметических способностей и способности к логическому рассуждению. Кроме того, оценивается запас фактических знаний ребенка; именно на это направлены вопросы субтеста «Осведомленность» шкалы Векслера. В целом, *IQ*-тесты для детей ориентированы на навыки, необходимые для успешного обучения в школе, — словарный запас, память, арифметические способности, решение задач. Поэтому неудивительно, что результаты выполнения этих тестов коррелируют с уровнем школьной успеваемости.

Стэнфорд—Бине и *WISC* обнаруживают сходство и по другим параметрам. В обоих тестах *IQ* ребенка рассматривается как функция от темпов его развития в сравнении с темпами развития других детей того же возраста. Это непосредственно отражено в формуле, по которой изначально рассчитывался *IQ* в тесте Стэнфорд—Бине: коэффициент интеллекта равен «умственному возрасту» (который определяется по тому, насколько трудными были задания, до которых дошел ребенок), деленному на «хронологический возраст» и умноженному на 100. По разным причинам эта формула больше не используется; теперь *IQ* вычисляется по отклонению показателя ребенка от среднего показателя в его возрастной группе. Дети, темп развития которых выше среднего, имеют *IQ* выше среднего; дети, темп развития которых ниже, чем в среднем по их возрастной группе, имеют и *IQ* ниже среднего. Таким образом, *IQ* ребенка — это показатель темпа развития. Кроме того, это по сути своей *относительный* показатель. Абсолютной метрики для измерения интеллекта ребенка, как, например, для измерения физических характеристик (высоты, веса и т. д.) не существует. *IQ* — это всегда результат сравнения ребенка с другими детьми.

Последнее, что имеют общее Стэнфорд—Бине и *WISC*, — это способ их предъявления. При предъявлении *IQ*-теста особое значение придается двум моментам. Первый — необходимость *стандартизации*. Как только что подчеркивалось, *IQ* ребенка — это относительная мера успешности выполнения им заданий в сравнении с успешностью выполнения этих заданий другими детьми. Единственный способ сделать показатель ребенка доступным для интерпретации — предъявлять тест и производить оценку одинаково для всех детей. Поэтому крайне важно, чтобы тестирующий знал и использовал стандартные инструкции. Второй момент — необходимость установления и поддержания *раппорта*. Предполагается, что *IQ* — это показатель оптимального уровня выполнения ребенком заданий, а оптимума мож-

но достичь только в том случае, если ребенок чувствует себя комфортно и мотивирован на тщательное обдумывание своих ответов. Лучше всего предъявляет тесты тот, кто достигает этих двух целей: придерживается стандартной процедуры и в то же время использует свои клинические навыки с тем, чтобы создать условия для оптимального выполнения ребенком заданий теста.

Примеры из табл. 11.3 дают возможность получить некоторое представление об IQ-тестах для дошкольников и школьников. Таблицы 11.4 и 11.5 дополняют картину интеллектуального развития тестами, предназначенными для других возрастных групп. В табл. 11.4 представлена подборка заданий из Шкал развития младенцев Бейли (Bailey, 1993), главного теста развития в младенчестве. Думаю, ни у кого не вызовет удивления тот факт, что тесты младенцев в значительно меньшей мере ориентированы на вербальные или академические навыки, чем тесты для детей старшей возрастной группы. В табл. 11.5 представлена подборка заданий, которые включены в тест Векслера, предназначенный для самого высокого уровня, — в Векслеровскую шкалу интеллекта взрослых (Wechsler, 1981). Этот тест — один из основных источников информации в исследованиях интеллекта лиц пожилого возраста, о которых пойдет речь в главе 13.

Таблица 11.4

Примеры процедур из шкал развития младенцев Бейли, второе издание

Возрастной уровень (в месяцах)	Измеряемая способность	Процедура	Зачет
1	Габитуация на звук погремушки	Встряхивайте погремушкой за головой ребенка через равные интервалы времени	Если ребенок демонстрирует ориентировочную реакцию, которая при повторях звукового сигнала угасает
6	Улыбается, видя собственное отражение в зеркале	Поставьте перед лицом ребенка зеркало	Если ребенок улыбается, увидев свое отражение в зеркале
12	Толкает машинку	Подталкивайте игрушечную машинку на глазах у ребенка; затем попросите его: «Толкай машинку, толкай машинку, как я»	Если ребенок целенаправленно толкает машинку так, что все ее колеса соприкасаются с поверхностью стола
17–19	Использует два разных слова	Фиксируйте спонтанное использование ребенком слов во время обследования	Если ребенок правильно использует (не имитируя) два слова
23–25	Указывает на пять картинок	Покажите 10 картинок обычных объектов (например, собака, книга, машина) и попросите: «Покажи мне _____»	Если ребенок правильно указывает или правильно называет по крайней мере пять картинок

Таблица 11.5

Типы заданий, включенных в Векслеровскую шкалу интеллекта взрослых (WAIS)

Осведомленность	Из чего состоит пар? Кто написал «Тома Сойера»?
Арифметический	Три женщины разделили между собой поровну 18 мячей для гольфа. Сколько мячей досталось каждой? Если 10 мужчин могут построить дом за 12 дней, сколько дней потребуется на постройку дома 4 мужчинам?
Субтест	Невербальная шкала
Недостающие части	Укажите на недостающую часть 

Источник: *Wechsler Adult Intelligence Scale-Revised*, D. Wechsler, 1981, New York: The Psychological Corporation.

Спорные вопросы и исследовательские парадигмы

Споры вокруг *IQ*-тестов, как правило, имеют отношение к использованию этих тестов в прагматических целях — например, для разделения школьников на потоки в соответствии с их показателями *IQ*. Здесь же основное внимание мы уделим теоретическим вопросам о развитии интеллекта. Два таких вопроса послужили толчком для многочисленных исследований и поводом для дискуссий: вопрос стабильности *IQ* и вопрос детерминант различий в *IQ*.

Тему стабильности можно рассмотреть вкратце, поскольку самое главное уже было сказано в главе 3, когда речь шла о лонгитюдном методе исследования. Для изучения стабильности требуется лонгитюдный подход, так как нас интересуют отношения между результатами ребенка на ранних этапах развития и результатами на дальнейших этапах развития. Особое значение имеет вопрос стабильности индивидуальных различий. Сохраняют ли дети свои относительные позиции по *IQ*-тестам в процессе развития; остаются ли высокими показатели тех, у кого они были высокими, и остаются ли низкими показатели тех, у кого они были низкими? Или возможны изменения? Обычно об этом судят по корреляциям между результатами первого и второго тестирования; чем выше коэффициент корреляции, тем выше стабильность. Поскольку *IQ*-тесты для детей созданы таким образом, что в каждой возрастной группе средний *IQ* один и тот же, появляется возможность оценить константность значения *IQ* самого по себе. *IQ* в конечном счете и есть относительный уровень, а константность относительного уровня подразумевает константность *IQ*. Мы, к примеру, можем задать вопрос, будет ли *IQ* ребенка, у

которого в 4-летнем возрасте IQ составлял 90 пунктов, в 6-, 10- или 20-летнем возрасте также равен 90.

IQ — одна из популярнейших тем лонгитюдных исследований в течение уже почти 70 лет. Подобные исследования могут и часто сталкиваются со всеми проблемами лонгитюдных исследований, которые рассматривались в главе 3. Испытуемые, которые имеют возможность и охотно соглашаются проходить многократные тестирования, не всегда представляют собой репрезентативную выборку, что ограничивает возможность генерализации полученных результатов на популяцию в целом. Выбывание в ходе исследования может оказаться избирательным, и в этом случае выборка станет еще более искаженной. Повторное выполнение одного и того же теста может привести к эффектам практики, а следовательно, к завышению результатов последних тестирований по сравнению с результатами первых тестирований. Кроме того, если в исследование включаются возрастные группы, для которых требуются разные IQ -тесты (например, младенцы и старшие дети), может стать актуальной проблема эквивалентности измерения.

Стоит отметить несколько выводов, которые можно сделать на основе исследований стабильности. Первая категория выводов касается прогностичности результатов, полученных в период младенчества. За исключением крайне низких показателей, успешность выполнения традиционных тестов для младенцев, таких как шкалы Бейли, обычно непрогностичны в отношении IQ на следующих этапах развития (Lipsitt, 1992). Этот факт известен уже давно. Он наводит на мысль об отсутствии непрерывности в развитии интеллектуальных навыков от младенчества к периоду детства: то есть то, что мы подразумеваем под словом «интеллект», у младенца, находящегося на сенсомоторной довербальной стадии, не то же самое, что у старшего ребенка или взрослого. Несомненно, в этом аргументе об отсутствии непрерывности есть доля истины; тем не менее последние исследования свидетельствуют о том, что этот аргумент имеет одно существенное ограничение. В ряде исследований было обнаружено, что *реакция на новизну* в младенчестве является *прогностичной* в отношении IQ — не абсолютно прогностичной, но с коэффициентом корреляции от 0,35 до 0,40 (Colombo, 1993; McCall & Mash, 1994). Таким образом, малыши, которых особенно интересует все новое и которые особенно чутко реагируют на новизну, в среднем имеют впоследствии самые высокие показатели IQ . Этот факт, в свою очередь, послужил стимулом для создания новой методики оценки интеллекта младенцев, которая строится на измерении реакции на новизну: теста интеллекта младенцев Фэгана (*The Fagan Test of Infant Intelligence*) (Fagan & Detterman, 1992; Fagan & Shepherd, 1986).

Что можно сказать о прогнозах, сделанных на основе показателей IQ в детском возрасте? Как только мы выходим за рамки младенчества, показатели интеллекта в разные возрастные периоды действительно начинают коррелировать между собой; однако стабильность все же далека от идеальной. В целом, коэффициент корреляции — а следовательно, и близость показателей IQ — тем выше, чем меньше разница между двумя сравниваемыми возрастами; она также тем выше, чем старше ребенок на момент первого тестирования: то есть стабильность IQ повышается с возрастом.

С момента создания первых IQ -тестов вокруг вопроса о причинах различий в IQ идут жаркие споры. Отчасти эти разногласия объясняются невозможностью

получить однозначные доказательства. Называют два основных источника различий в *IQ*: разницу в генах, с которыми рождается человек, и разницу в среде, где он воспитывается. Достаточно легко можно спланировать исследование, в котором эти два фактора были бы разделены; все, что необходимо, это, поддерживая неизменным один фактор, систематически изменять другой. Точно так же легко догадаться, что проведение таких исследований невозможно. Поэтому мы вынуждены довольствоваться данными, не отвечающими всем нашим требованиям. При противопоставлении факторов наследственности и факторов среды особое значение приобретают результаты исследований двух типов: исследований близнецов и исследований приемных детей.

В главе 10, говоря о темпераменте, мы уже упоминали о близнецовом методе. Исследования близнецов строятся на факте существования двух типов близнецовых пар.Monoзиготные, или однояйцевые близнецы развиваются из одной яйцеклетки и поэтому идентичны; дизиготные, или двужайцевые близнецы появляются из разных яйцеклеток и поэтому общность их генов составляет около 50 %, то есть такая же, как и у обычных сиблингов. Таким образом, мы имеем дело с естественным экспериментом, в котором изменяется генетическая переменная. Если гены влияют на *IQ*, то *IQ* монозиготных близнецов должны быть ближе, чем *IQ* дизиготных близнецов. И реальная ситуация как раз такова. Исследователи сообщают о коэффициенте корреляции между *IQ* монозиготных близнецовых пар около 0,8; коэффициент же корреляции у дизиготных пар обычно в пределах 0,5–0,6 (Nichols, 1978; Plomin, 1990).

Генетические интерпретации результатов близнецовых исследований имеют очевидные недостатки. Возможно, в среднем, монозиготные близнецы воспитываются в более сходных условиях, чем дизиготные близнецы. В конечном счете, по внешнему виду и поведению монозиготные близнецы в некоторых отношениях больше похожи друг на друга, чем дизиготные близнецы. Эту проблему можно было бы разрешить, проведя исследование монозиготных близнецов, разлученных в раннем детстве и воспитывавшихся в разной среде. Найти такие пары непросто — подобных исследований осуществлено совсем немного, и ни в одном из них не используется достаточно большая выборка. Кроме того, близнецов нельзя разлучить в научных целях; они воспитываются порознь по разным причинам, при разных обстоятельствах, а такой недостаток контроля затрудняет интерпретацию. Эту ситуацию осложняет и тот факт, что многие из первых исследований разлученных близнецов страдают рядом методологических погрешностей, и поэтому их результаты трудно интерпретировать (Kamin, 1974). Тем не менее данные, полученные в ходе таких исследований — и не только первых, но и последних, более совершенных с методологической точки зрения работ (Bouchard, Lykken, McGue, Segal, & Tellegen, 1990), — представляют собой довольно убедительное доказательство генетической модели. Обнаружено, что коэффициент корреляции между *IQ* монозиготных близнецов, воспитывающихся порознь, в среднем составляет 0,75 — лишь немного меньше, чем коэффициент корреляции у монозиготных близнецов, воспитывающихся вместе, и выше, чем коэффициент корреляции у воспитывающихся вместе дизиготных близнецов.

Исследованиям приемных детей посвящено больше работ, чем исследованиям близнецов. Точкой отсчета для таких исследований является тот факт, что коэффициент корреляции между *IQ* родителей и *IQ* детей составляет около 0,5. То есть дети, как правило, по интеллекту похожи на родителей; вопрос — почему? От каждого из родителей ребенок получает по 50 % генов; то есть корреляция может быть обусловлена влиянием генетических факторов. Однако каждый из родителей представляет собой важнейшую часть социальной среды, окружающей ребенка; то есть корреляция может быть обусловлена влиянием и средовых факторов. Исследования приемных детей позволяют разделить эти два типа факторов. Мы получаем возможность проанализировать корреляции двух видов. Первый — корреляция между *IQ* приемных детей и *IQ* приемных родителей. В этом случае задействуются факторы среды, а влияние генетики исключается. Второй вид — корреляция между *IQ* приемного ребенка и *IQ* его биологических родителей. В этом случае исключается воздействие уже факторов среды (если не считать влияние среды в пренатальном и, возможно, раннем постнатальном периоде); в то же время задействуются генетические факторы.

Исследования приемных детей позволяют сделать два основных вывода (Scarr & Weinberg, 1983; Turkheimer, 1991). Первый — *IQ* ребенка коррелирует с *IQ* биологических родителей в большей мере, чем с *IQ* приемных родителей. Это подтверждает значимость генетических факторов. Второе, что удалось обнаружить, — это то, что *IQ* приемных детей обычно выше среднего, и обычно выше среднего *IQ* их биологических родителей. Поскольку корреляция — это показатель относительного уровня, такая средняя разница может иметь место даже в том случае, когда коэффициент корреляции между показателями *IQ* детей и их биологических родителей достаточно высок. Часто разница между родителями и детьми объясняется регрессией к среднему, феноменом, который обнаруживается и при сравнении поколений (то есть дети родителей, имеющих *IQ* ниже среднего, как правило, превосходят своих родителей). Однако часть разницы почти наверняка является отражением благотворного воздействия среды в приемной семье. Эти семьи, как правило, имеют множество преимуществ перед семьями биологических родителей; и они явно стимулируют интеллектуальное развитие детей, которые в них воспитываются. Таким образом, результаты исследований приемных детей свидетельствуют в пользу и генетических, и средовых факторов.

Необходимо отметить, что исследования приемных детей, как и близнецовые исследования, имеют некоторые ограничения. Об указанных выше корреляциях можно говорить только при наличии показателей *IQ* во всех задействованных группах. Поскольку в некоторых исследованиях получить информацию об *IQ* части выборки не удалось, об *IQ* судили по образовательному уровню — сомнительная и потенциально искажающая результаты процедура. Недоверие к результатам вызывает и возможность избирательного помещения детей в приемные семьи — то есть стремление органов, занимающихся вопросами усыновления, подбирать приемные семьи в соответствии с характеристиками биологических родителей. В той мере, в какой эта избирательность действительно существует, она сильно затрудняет любые интерпретации коэффициентов корреляции между *IQ* родителей и

детей. Еще одна проблема — незначительное разнообразие приемных семей. Как указывалось выше, эти семьи не являются случайной выборкой из популяции; как правило, по многим показателям они поднимаются над средним уровнем. Это означает, что приемные семьи представляют собой довольно однородную группу, а такая однородность затрудняет корреляционный анализ. Чем меньше изменчивость переменной (в нашем случае, характеристик приемных семей, включая и IQ приемных родителей), тем меньше вероятность того, что эта переменная будет значимо коррелировать с другими. Этот фактор накладывает серьезные ограничения на возможности интерпретации низкого коэффициента корреляции между интеллектом приемных родителей и приемного ребенка.

Память

В следующих разделах этой главы мы перейдем от общих подходов к рассмотрению специфической проблематики исследований когнитивного развития. Из всех тем, заслуживающих внимания, я выделяю две, особенно интересные и информативные. Первая тема имеет давнюю историю. Память является предметом психологического изучения с момента зарождения психологии как науки; а память детей всегда вызывала интерес и родителей, и психологов. Изучение второй темы началось не так давно. Хотя ее предпосылки оформились в предшествующие годы, такая область исследования психологии развития, как «теория сознания» (*theory of mind*) стала интенсивно разрабатываться лишь в последние 8–10 лет.

Память в младенчестве

При рассмотрении исследований младенцев в главе 10 мы не затрагивали тему памяти. Тем не менее многие из описанных там процедур, что бы ни измерялось с их помощью, так или иначе позволяют получить представление и о памяти младенца. Работы Пиаже, к примеру, свидетельствуют о том, что дети способны запоминать информацию и действовать в соответствии с ней уже на самых ранних этапах развития. Аналогично, многие проявления привязанности, такие как предпочтение ребенком матери незнакомому человеку, очевидно, подразумевают работу памяти.

Главные процедуры исследования памяти младенцев основаны на двух подходах, описанных в главе 10: на парадигме «габитуация/дегабитуация» и на методе предпочтений Фанца. Запоминание — неотъемлемая часть процессов габитуации и дегабитуации. Ребенок может привыкнуть к стимулу, лишь сохраняя информацию об этом стимуле и узнавая его при повторном воздействии. Если бы не было памяти на прошедшие события, не могло бы быть и габитуации. Аналогично, младенец может продемонстрировать дегабитуацию на новый стимул, только если помнит исходный стимул и понимает, что новый стимул от него чем-то отличается. Как отмечалось в главе 10, габитуация и дегабитуация наблюдаются уже у новорожденных, что доказывает наличие определенных мнемических возможностей с самого рождения.

Способность к запоминанию не является неотъемлемой частью метода предпочтений, но ее легко можно включить в эту процедуру. Все, что для этого необходи-

мо, это сделать главным параметром, по которому будут различаться два стимула, то, насколько они знакомы ребенку. Предположим, что в серии попыток одним из стимулов неизменно является треугольник; другой стимул изменяется. Вопрос состоит в том, начнет ли младенец с каждой попыткой смотреть на относительно незнакомый стимул дольше, чем на относительно знакомый. Дети старше 2-месячного возраста действительно демонстрируют предпочтение новизны; дети младше 2-месячного возраста могут предпочитать знакомые стимулы (например, Wetherford & Cohen, 1973). Для нас важно, что предпочтение и нового и знакомого подразумевает работу памяти.

Возможности описанных выше методов не ограничиваются простым доказательством наличия памяти; их можно использовать для изучения системы памяти в раннем возрасте. Можно, например, измерить длительность сохранения информации в памяти младенца, изменяя временной интервал между предъявлением стимула и проверкой того, помнит ли его ребенок. Или же, можно узнать о том, какие аспекты стимула замечает и запоминает младенец, посмотрев, на изменение каких аспектов стимула он реагирует в фазу дегабитуации. Если, к примеру, стимул состоит из компонентов *A*, *B* и *C*, можно осуществить проверку с помощью процедуры габитуации/дегабитуации, изменяя только *A*, только *B* или только *C*. В исследованиях подобного рода были выявлены довольно впечатляющие возможности младенцев, а также важнейшие направления, в которых происходит совершенствование памяти в первые годы жизни (Olson & Sherman, 1983; Rovee-Collier, 1987).

Для изучения памяти в раннем возрасте используются и другие процедуры. Рассмотрим приспособление, изображенное на рис. 11.2. Как явствует из рисунка, лента, соединяющая щиколотку малыша и подвесные игрушки, дает возможность ребенку их раскачивать. Уже в 2-месячном возрасте малыши способны установить наличие этой связи (форма оперантного обучения). Как только реакция закрепилась, можно по-разному изменять ситуацию, проверяя память младенца. Мы можем оценить время хранения информации в памяти, повторно подвесив игрушки через некоторое время и посмотрев, будет ли малыш двигать ногой. Можно оценить специфичность запоминаемой информации, подвешивая новые игрушки, имеющие разную степень сходства с игрушками, использовавшимися для обучения. В этих исследованиях также обнаруживаются впечатляющие способности детей раннего возраста и совершенствование этих способностей в период младенчества (Rovee-Collier & Bhatt, 1993).

Мнемический процесс, который наглядно демонстрируют все описанные выше исследования, называется **узнаванием**. Психологическое определение «узнавания» ничем не отличается от повседневного, словарного определения: это осознание того, что нечто новое является таким же, как то, что встречалось ранее. Таким образом, младенец, у которого наблюдается габитуация, узнает, равно как и младенец, реагирующий на знакомое лицо матери. Узнавание можно сопоставить с другим базовым мнемическим процессом — **воспроизведением**. Воспроизведение подразумевает активное восстановление в памяти того, что *не* присутствует непосредственно в данный момент. Ребенок, рассказывающий о том, как прошел неделю назад его день рождения, демонстрирует воспроизведение; равно как и ребенок, изображающий на рисунке праздник.

Помимо того что способность к воспроизведению, скорее всего, отсутствует при рождении и абсолютно точно наличествует к концу периода младенчества, об этой способности у младенцев практически ничего не известно. Воспроизведение трудно изучать, поскольку методы, используемые со старшими детьми, такие как вербальные отчеты или рисунки, к малышам неприменимы. Фактически нет единого мнения даже о том, какие именно действия младенца могли бы недвусмысленно указывать на его способность к воспроизведению. Тем не менее убедительно доказано, что некоторые формы воспроизведения могут наблюдаться к концу первого года жизни. Одно из доказательств — поиск младенцем предметов, удаленных из поля зрения. К 9–10-месячному возрасту большинство младенцев способны искать предметы, полностью исчезнувшие из поля зрения, что, согласно Пиаже, соответствует 4 стадии в 6-ступенчатой модели развития представлений о постоянстве объекта. К этому возрасту большинство детей запоминает местоположение знакомых вещей, например, кухонного шкафа, в котором хранятся любимые кукурузные хлопья (Ashmead & Perlmutter, 1980). Поиск объекта, который отсутствовал в поле зрения несколько дней, несомненно, подразумевает определенную способность к воспроизведению.

Исследования отсроченного подражания — то есть способности воспроизводить некую модель поведения, виденную в прошлом, — еще один источник данных о памяти. Пиаже был первым, кто начал изучать отсроченное подражание; результаты его исследований свидетельствуют о появлении подражания в возрасте 18–24 месяцев. Это согласовывалось с его гипотезой о том, что необходимым предпосылкой способности к отсроченному подражанию является способность к репрезентации и воспроизведению, которая формируется не раньше, чем к концу периода младенчества. Ученые, проводившие исследования после Пиаже, согласились с его теоретическим обоснованием, однако пришли к иному выводу относительно сроков. Так, было показано, что 9-месячные дети способны подражать новым действиям (например, нажимают на кнопку, в результате чего раздается звук), свидетелями которых они были за 24 часа до этого (Meltzoff, 1988b). К 14-месячному возрасту период сохранения информации в памяти увеличивается до целой недели (Meltzoff, 1988a).

Младенцы могут имитировать и запоминать не только отдельные действия, но и простейшие цепочки действий. Одиннадцатимесячные малыши, к примеру, могут повторить показанную им взрослым последовательность, состоящую из двух действий (например, сделать погремушку: положить в коробку пуговицу, а затем потрясти коробку); 13-месячные дети могут освоить уже трехкомпонентную последовательность (например, сделать более сложную погремушку: положить шарик в одну полусферу, соединить две полусферы, вставив одну в другую, а затем потрясти погремушкой. См.: Bauer & Mandler, 1992). К 20-месячному возрасту младенцы могут воспроизводить последовательности действий не только незамедлительно после показа, но и с 2-недельной отсрочкой (Bauer & Mandler, 1992). В одном из лонгитюдных исследований были получены данные о том, что некоторые малыши, прошедшие первое тестирование в 11-месячном возрасте, смогли воспроизвести некоторые аспекты полученного ими опыта по прошествии целого года (McDonough & Mandler, 1989)!

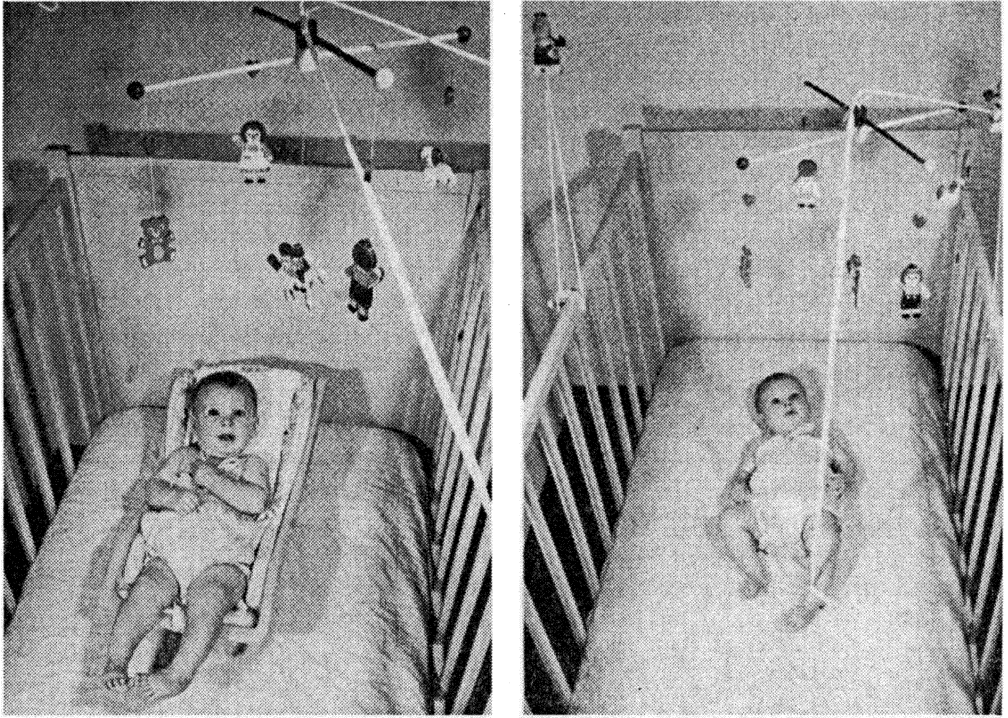


Рис. 11.2. Экспериментальное приспособление для изучения способности младенцев к научению и запоминанию. Привязанная к щиколотке малыша лента (как на фото справа) позволяет ему раскачивать игрушки, подвешенные над кроваткой. Научение демонстрируется интенсивными движениями ног в ситуации, когда к ноге привязана лента, а над кроватью подвешены игрушки. (Фотографии предоставлены Dr. C. K. Rovee-Collier)

Как только мы переходим от младенчества к другим возрастным периодам, количество исследований процесса воспроизведения начинает существенно превышать количество исследований процесса узнавания. Давно известно, что старшие дети, в среднем, воспроизводят информацию лучше, чем младшие. Эта закономерность проявляется в различных контекстах и при различных формах воспроизведения, включая воспроизведение более сложных последовательностей, способность к которому, как мы только что отметили, формируется в период младенчества. Особый интерес представляет вопрос о том, чем обусловлено это совершенствование памяти. Обратимся к рассмотрению двух видов исследований, в которых осуществляется попытка выявить то, что является базисом для развития памяти.

Мнемонические стратегии

Главная идея, лежащая в основе изучения **мнемонических стратегий**, заключается в том, что улучшение способности к воспроизведению в ходе развития является результатом не только — и возможно не столько — простого увеличения объема памяти. Этот прогресс отражает процесс освоения неких *действий* — неких мнemo-

нических стратегий, — помогающих запоминать, сохранять и воспроизводить определенную информацию. Такие стратегии могут использоваться при предъявлении материала для запоминания, в период отсрочки между предъявлением и проверкой или при попытке восстановить материал в памяти. Эти стратегии могут принимать разнообразные формы: одни из них — вербальные, другие — невербальные; одни — простые, другие — очень сложные. Их общее свойство состоит в том, что они облегчают работу памяти.

Таблица 11.6

Примеры мнемонических стратегий

Общая стратегия	Экспериментальное задание	Приемы, используемые ребенком
Вербальное повторение	По очереди предъявляются и убираются 10 картинок с изображениями знакомых, но не связанных между собой предметов. Задача ребенка — вспомнить и назвать как можно больше из показанных картинок	Сингулярное повторение: ребенок повторяет название каждой из картинок во время ее предъявления — например: «яблоко, яблоко, яблоко, флаг, флаг, флаг...» Кумулятивное повторение: ребенок называет каждую картинку во время ее показа, а затем повторяет название всех показанных картинок — например: «яблоко, яблоко-флаг, яблоко-флаг-луна...»
Переработка	Предъявляется 20 пар знакомых, но не связанных между собой слов (например, «корова—галстук», «машина—дерево»). В дальнейшем предъявляется только один из элементов каждой пары; задача ребенка — вспомнить и назвать второй элемент	Образная переработка: ребенок создает мысленный образ, так или иначе связывающий элементы каждой из пар — например: «корова носит галстук, машина заезжает на дерево...» Вербальная переработка: ребенок создает предложения или фразы, так или иначе объединяющие элементы каждой пары — например: «корова носила галстук», «машина заехала на дерево...»
Упорядочение	Предъявляется 20 картинок с изображениями знакомых предметов, которые относятся к 4 понятийным категориям: «животные», «транспорт», «одежда» и «мебель». При предъявлении картинки не группируются по категориям. Задача ребенка — вспомнить и назвать как можно больше картинок	Группирование: ребенок упорядочивает и воспроизводит названия картинок по категориям, к которым они относятся, — например, «собака, лошадь, верблюд, медведь, белка, автомобиль, грузовик, самолет...»

Прежде чем обратиться к вопросу о том, как изучаются мнемонические стратегии, рассмотрим вкратце, что они из себя представляют. В табл. 11.6 приведены примеры трех категорий мнемонических стратегий. Это отнюдь не единственные стратегии, которые могут использовать дети; в действительности, мы уже приводили пример другой мнемонической стратегии — стратегии адекватного распределения времени запоминания (Dufresne & Kobasigava, 1989), описанного в главе 2. Тем не менее стратегии повторения, упорядочения и переработки привлекают наибольшее внимание исследователей.

Как изучать мнемонические стратегии? Некоторые стратегии по своей сути имеют внешние проявления, и поэтому их можно легко исследовать. Ведение записей, к примеру, довольно распространенная и легко наблюдаемая мнемоническая стратегия. Сюда же можно отнести обращение к родителям с просьбой помочь что-то запомнить. Наверное, по причине того, что такие стратегии, имеющие внешние проявления, кажутся столь очевидными, у психологов они не вызывают большого интереса. Интерес представляют скорее внутренние, мысленные стратегии, как те, которые приведены в таблице, — стратегии, необходимость в использовании которых возникает тогда, когда такие «вспомогательные средства», как записи или родители недоступны. Как выявить мнемоническую стратегию?

Один из способов — это *логический вывод* об использовании стратегии на основе успешности выполнения испытуемым теста памяти. Предположим, что у нас сложный в вербальном отношении тест памяти, что мы изучаем вербально зрелых старших детей и вербально незрелых младших детей и что результаты старшей группы оказались выше результатов младшей группы. Разумный (хотя, конечно, и не бесспорный) вывод из этого заключается в том, что старшие дети лучше справились с заданием, поскольку использовали свои вербальные навыки. Возможно, к примеру, что в период отсрочки они в отличие от младших детей проговаривали то, что нужно было запомнить. Этот вывод приобретет большую убедительность, если мы выявим изменения не только уровня, но и *модели* выполнения задания. Например, классический результат исследования памяти взрослых — так называемый «эффект первичности», означающий тенденцию лучше запоминать первые элементы ряда. Обычно этот эффект объясняют вербальным повторением первых элементов. В аналогичных сравнительных исследованиях было обнаружено, что у старших детей наблюдается эффект первичности, а у младших, как правило, нет (например, Cole, Frankel, & Sharp, 1971). Эта модель согласуется с гипотезой о том, что с возрастом повышается значимость вербального повторения.

Второй подход — *вызвать* использование стратегии. В этом случае мы не строим предположений относительно наличия или отсутствия стратегии; мы просто даем детям инструкцию использовать определенную стратегию, а затем наблюдаем эффект. Можно, к примеру, попросить половину испытуемых повторять материал для запоминания в период отсрочки, а второй, контрольной группе таких инструкций не давать. Исследования вызванного использования стратегий позволяют сделать несколько выводов. Основной из них заключается в том, что подобные инструкции оказывают благотворное воздействие; дети, которым подсказывают использовать определенную стратегию, обычно лучше справляются с заданием.

Заметьте, что здесь мы имеем дело с конвергенцией данных, о которой говорилось в главе 6: исследования, в которых делается логический вывод об использовании стратегии, свидетельствуют о том, что используется некая мнемоническая стратегия, имеющая благотворный эффект; исследования использования стратегии по подсказке доказывают, что эта стратегия действительно производит благотворный эффект. В подобных исследованиях можно не ограничиваться простым сопоставлением «использование стратегии/не использование стратегии», а анализировать эффекты разного рода стратегий. Например, было продемонстрировано, что кумулятивное повторение (повторение названий всех предъявляемых объектов) эффективнее некумулятивного (повторения названия последнего из предъявленных объектов) (Ferguson & Bray, 1976). Аналогично, семантическая категоризация объектов обычно дает лучшие результаты, чем структурирование материала по перцептивным или фонологическим характеристикам (Cox, Ornstein, Naus, Maxfield, & Zimler, 1989).

Хотя описанные выше исследования весьма информативны, они имеют известные ограничения. Исследования, в которых делается логический вывод об использовании определенной мнемонической стратегии, имеют очевидный недостаток, заключающийся в том, что об использовании стратегии судят по косвенным признакам; мы не знаем наверняка, что именно делает ребенок. Исследования вызванного использования стратегий устраняют этот недостаток, однако добавляют новый: поскольку мы заставляем ребенка использовать некую стратегию, нам неизвестно, что бы он делал без подсказки.

Прорыв в исследованиях памяти детей произошел в 1960-х годах и связан с разработкой экспериментальных ситуаций, в которых дети спонтанно создавали мнемонические стратегии, имеющие по крайней мере некоторые внешние, а следовательно, измеримые проявления. Это открытие дало возможность соединить лучшее, что есть в двух других подходах: изучать стратегии, использование которых одновременно и спонтанно, и выражено внешне. Поскольку пионером в этой области был Джон Флэйвелл, приведу в качестве примера одно из его ранних исследований. Подчеркнем, однако, что этот пример — лишь один из множества; исследовательская группа Флэйвелла (например, Flavell, Friedrichs, & Hoyt, 1970) и другие исследователи (например, Bjorklund, Coyle, & Gaultney, 1992; Miller, 1990) проявили определенную изобретательность при разработке экспериментальных условий, позволяющих непосредственно наблюдать использование стратегий.

Дошкольникам, 2- и 5-классникам Флэйвелл, Бич и Чински (Flavell, Beach & Chinsky, 1966) предъявляли картинки семи знакомых предметов. Экспериментатор указывал на картинки в определенном порядке; задачей детей было воспроизвести названия картинок в правильном порядке. Показ картинок и проверку разделяла 15-секундная отсрочка. Во время проведения исследования на ребенке был игрушечный гермошлем, а на период отсрочки отверстие для глаз закрывалось, чтобы ребенок не смотрел на картинки. Щиток для глаз выполнял и еще одну функцию, позволяя экспериментаторам следить за губами ребенка во время отсрочки. В задачу экспериментатора входила запись любых вербализаций ребенка, как открытых, так и полуоткрытых (перед проведением исследования экспериментаторы обучались чтению по губам). Особый интерес, конечно же, представляли слу-

чаи повторения. Дети действительно повторяли, однако вероятность использования этой стратегии зависела от возраста: заметное повторение демонстрировали 17 из 20 5-классников, в сравнении лишь с 2 из 20 дошкольников. При последующем использовании этой процедуры (Keene, Cannizzo, & Flavell, 1967) было обнаружено, что внутригрупповые индивидуальные различия (среди первоклассников) в использовании стратегии повторения коррелируют с успешностью воспроизведения — то есть дети, спонтанно повторявшие запоминаемый материал, лучше справились с заданием. Типичные выводы из исследований, посвященных памяти, — указание на тенденцию использовать определенные стратегии и на благотворный эффект их использования.

В исследованиях, подобных тем, которые были описаны выше, детям младше 5–6 лет часто не удается изобрести мнемоническую стратегию. Чтобы не возникло впечатление, что это абсолютное правило, рассмотрим исследование Уэлман, Риттер и Флэйвелла (Wellman, Ritter, & Flavell, 1975). Их испытуемыми были 3-летние дети, а экспериментальная процедура кратко изложена в следующем отрывке:

Я расскажу тебе историю о собачке. Смотри, вот она на лужайке [поверхность стола]. Она любит играть, бегают, прыгает... Но она так много играла, что проголодалась. И вот она пошла поискать что-нибудь поесть. Когда она искала, что бы ей поесть, то проходила мимо этой конуры, этой конуры, этой конуры и этой конуры [собаку проносят мимо четырех чашек]. А потом, чтобы найти что-нибудь поесть, она забежала в эту конуру [собаку прячут]. Знаешь, у меня есть еще одна игрушка; она помогла бы рассказать эту историю. Я пойду, возьму ее, она нам понадобится (p. 781).

С этого момента инструкции для разных экспериментальных условий расходятся. Детей из группы «Ожидание» просто просили подождать, пока принесут игрушку. Детей из группы «Запоминание» просили запомнить, где оказалась собака. Вопрос, естественно, состоял в том, будет ли поведение детей, которых просили запомнить, где осталась собака, отличаться от поведения тех, кого просили просто подождать. Ответ состоит в том, что дети из группы «Запоминание» действительно вели себя иначе, и весьма разумно: в период отсрочки они значительно чаще неотрывно смотрели на нужную чашку или прикасались к ней. Да, это очень простые мнемонические стратегии, но это *стратегии*, и они доступны 3-летним детям. В других исследованиях с использованием подобной процедуры «игры в прятки» были обнаружены зачатки стратегического поведения уже в 18-месячном возрасте (DeLoache, Cassidy, & Brown, 1985). Здесь мы имеем еще один пример того, с чем периодически сталкиваются специалисты в области психологии развития: не следует делать поспешных выводов о том, что у маленького ребенка полностью отсутствует определенная способность (в нашем случае способность к использованию мнемонических стратегий).

Добавим также, что не следует делать поспешных выводов и о том, что некоторая способность на каком-то этапе достигает полного развития. Развитие мнемонических стратегий не завершается в 10–11 лет, хотя к этому возрасту дети весьма умело используют такую простую стратегию, как вербальное повторение. В последние годы исследователи проявляют все больший интерес к сложным мнемоническим стратегиям, освоение которых происходит не раньше подросткового возраста или даже взрослости. Наверное, особого внимания заслуживают работы Энн

Браун и ее соавторов, изучавших стратегии, применяемые к школьному материалу (Brown, 1982). Общий обзор и рассмотрение некоторых стратегий, которые осваиваются в зрелом возрасте, можно найти в работе Pressley, Levin, & Bryant, 1983.

Конструктивная память

Какое бы значение ни имели стратегии, их использование не объясняет все мнемические феномены или все важнейшие мнемические изменения, происходящие в процессе развития. Рассмотрим некоторые недостатки исследований стратегий памяти. В этих исследованиях обычно используется бессмысленный материал (например, списки не связанных между собой слов); однако ясно, что в реальной жизни мы запоминаем преимущественно то, что имеет смысл. В исследованиях стратегической памяти все внимание уделяется произвольным мнемическим процессам; однако часто (и возможно, в большинстве случаев) в реальной жизни мы используем память непроизвольно или нецеленаправленно, в том смысле, что мы запоминаем и помним то, что никогда не пытались запомнить специально. Наконец, исследования стратегической памяти подразумевают изучение самостоятельных, не связанных между собой способов хранения или восстановления информации в памяти. Однако не всю когнитивную деятельность, связанную с памятью, можно объяснить с точки зрения целенаправленного использования стратегий, которые могут рассматриваться по отдельности.

Исследования **конструктивной памяти** направлены на изучение того, как влияет на память система общих знаний. В основе этого подхода лежит мысль о том, что память — это форма практического познания, форма познания, отвечающая за хранение и реконструкцию информации, полученной в прошлом. Как любая форма познания, память подразумевает действие и понимание, а не просто пассивную регистрацию входящего сигнала. И как в любой форме познания, в ней происходят изменения, параллельные изменениям в понимании окружающего мира: «Знания человека сильно влияют на то, что он узнает, запоминает и помнит. Однако в процессе развития знания человека сильно изменяются, и эти изменения стимулируют изменения мнемического поведения» (Flavell et al., 1993, p. 255).

Рассмотрим несколько примеров. Пэрис и Картер (Paris & Carter, 1973) предъявляли для запоминания 7- и 10-летним детям список из 21 предложения. Эти предложения были сгруппированы в семь, не связанных между собой «рассказов», каждый из которых состоял из трех предложений. Один из рассказов, к примеру, звучал следующим образом: «Птица сидит в клетке. Клетка стоит под столом. У птицы желтое оперенье». Данной модели соответствовали все семь рассказов, состоящих из трех предложений каждый: две посылки, первая — устанавливающая отношения между *A* и *B* (например, между птицей и клеткой) и вторая — устанавливающая отношения между *B* и *C* (например, между клеткой и столом), плюс утверждение, логически не связанное с двумя предыдущими (например, «У птицы желтое оперенье»). Затем дети выполняли тест на узнавание предложений. Тест был составлен из четырех предложений на каждый из семи «рассказов». Одно предложение было истинной посылкой, например: «Птица сидит в клетке». Второй была слегка измененная ложная посылка, например: «Клетка стоит на столе». Третьим был истинный вывод об отношениях между *A* и *C*, например: «Птица сидит

под столом». Четвертым был ложный вывод об отношениях между А и С, например: «Птица сидит на столе». Таким образом, только одно из четырех предложений действительно входило в предъявлявшийся ранее список; однако второе предложение логически вытекало из информации, данной в списке.

У 7- и 10-летних детей наблюдались ошибки одного и того же типа. Дети из обеих групп реже всего ошибались в отношении трех из четырех предложений: истинной посылки, ложной посылки и ложного вывода. Однако в отношении истинных выводов ошибки совершались примерно в 60–70 % случаев. То есть дети демонстрировали устойчивую тенденцию «узнавать» предложения, которых до этого не видели. Подобные ответы рассматриваются как ошибки, но заметьте, ошибки интеллектуальные. Что делает ребенок, так это интерпретирует информацию, которую дают две посылки, и приходит, таким образом, к конечному заключению. В большинстве ситуаций ребенок осмысливает информацию так, как (за исключением ситуации выполнения лабораторных тестов на узнавание) это было бы вполне разумно.

Что делают дети, столкнувшись с более объемным материалом, чем «рассказы», использовавшиеся Пэрис и Картер? В исследовании Пэрис (Paris, 1975) дошкольникам и учащимся 1–5 классов читали следующую историю:

Линда играла с новой куклой во дворе своего большого красного дома. Внезапно она услышала странный звук, доносившийся из-под крыльца. Это было хлопанье крыльев. Она очень хотела помочь, но не знала, что делать. Она забежала в дом и взяла в кладовой коробку из-под обуви. Затем Линда порылась в своем письменном столе и нашла восемь листов желтой бумаги. Она разрежала ее на мелкие кусочки и разложила их на дне коробки. Линда осторожно подняла беспомощное существо и взяла к себе. Ее учительница знает, что нужно делать (р. 233).

Прослушав эту историю, дети должны были ответить на восемь вопросов:

- Была ли кукла Линды новой?
- Взяла ли Линда коробок от спичек?
- Раздавался ли звук из-под крыльца?
- Играла ли Линда за домом?
- Нравилось ли Линде заботиться о животных?
- Отнесла ли Линда найденное в полицию?
- Линда нашла лягушку?
- Использовала ли Линда ножницы? (р. 233)

Обратите внимание на кардинальное различие первых четырех вопросов и последних. Первые касаются информации, непосредственно данной в тексте. А на последние можно ответить, только сделав логические выводы из того, что было сказано. Главное — что обнаруживается не только в исследованиях Пэрис, но и в других исследованиях подобного рода — это то, что даже маленькие дети приходят к логическим заключениям, необходимым для того, чтобы ответить на вторую половину вопросов. Склонность делать логические выводы, а также максимальный уровень сложности этих выводов возрастают в процессе развития (Paris & Upton, 1976). Однако уже с раннего детства — в рассказе ли, в разговоре, в увиденных событиях, в чем угодно — запоминается смысл, а не формальные подробности.

Рассмотрим последний пример на эту тему. Мы уже отмечали, что даже младенцы обладают определенной способностью запоминать, в каком порядке обычно следуют события. Обладают такой способностью, конечно же, и старшие дети, запоминающие более длинные и сложные последовательности. Зафиксированная в памяти информация о порядке и структуре знакомых событий называется **сценарием**. В процессе развития в сознании детей формируется ряд сценариев. Исследуются такие сценарии, как сценарий купания, выпекания печенья, походов в «Макдональдс», празднования дня рождения. Сценарии хранятся в памяти и влияют на процесс запоминания другой информации. Дети (а в действительности, любой из нас) осмысливают новый опыт, связывая его со знакомыми сценариями. В целом, дети запоминают то, что соответствует структуре знакомого сценария, лучше, чем то, что ее нарушает; а также запоминают события, занимающие центральное место в сценарии, лучше, чем события периферические (Mandler, 1983; McCartney & Nelson, 1981). Дети могут даже исказить информацию при воспроизведении, приводя ее в соответствие с неким сценарием. В одном из исследований, к примеру, дошкольники, слышавшие фразу «дети принесли подарки» в конце рассказа о праздновании дня рождения, были склонны при пересказе ставить ее в начале (Hudson & Nelson, 1983). Мы еще столкнемся с подобной конструктивной «корректировкой» в главе 12 в контексте рассмотрения вопроса о том, как представления детей о различиях между полами влияют на то, что они запоминают о поведении мальчиков и девочек или мужчин и женщин.

В этом разделе мы лишь затронули тему воздействия представлений детей на их память. Более подробно она анализируется в работах Bjorklund, 1987 и Flavell, 1993.

Теория психики

Методы изучения

Рассмотрим простой сюжет, который иллюстрирует рис. 11.3. Для любого взрослого ответ на вопрос, где Салли будет искать свой мяч, очевиден — в корзине, где она его оставила. У нее не было никакой возможности узнать, что в ее отсутствие мяч переложили в другое место. Однако оказывается, что для многих детей дошкольного возраста это не так очевидно. Большинство детей заявляют, что Салли заглянет в коробку. Они отвечают с точки зрения того, что им самим известно о ситуации, а не с точки зрения того, о чем, скорее всего, могла бы подумать Салли.

Задание «Салли/Энн» — пример методики, используемой в новой и весьма интересной области исследования, области, которую называют **теорией сознания**. Теория психики имеет отношение к пониманию психического мира, к тому, что думают дети о таких феноменах как желания, намерения, эмоции и (как в примере) мнения. Интерес к подобным вопросам, естественно, возник уже давно. В частности, психическим феноменам уделяется внимание в пиажетианском подходе. Рассмотренные в этой главе исследования принятия чужой точки зрения фактически — предшественники современных работ, посвященных теории сознания. Однако сегодня круг охвата изучаемой проблематики значительно шире, чем у Пиаже, делавшего лишь первые шаги в данном направлении; за последнее время родилась масса теорий, спорных вопросов, было сделано множество открытий. Это

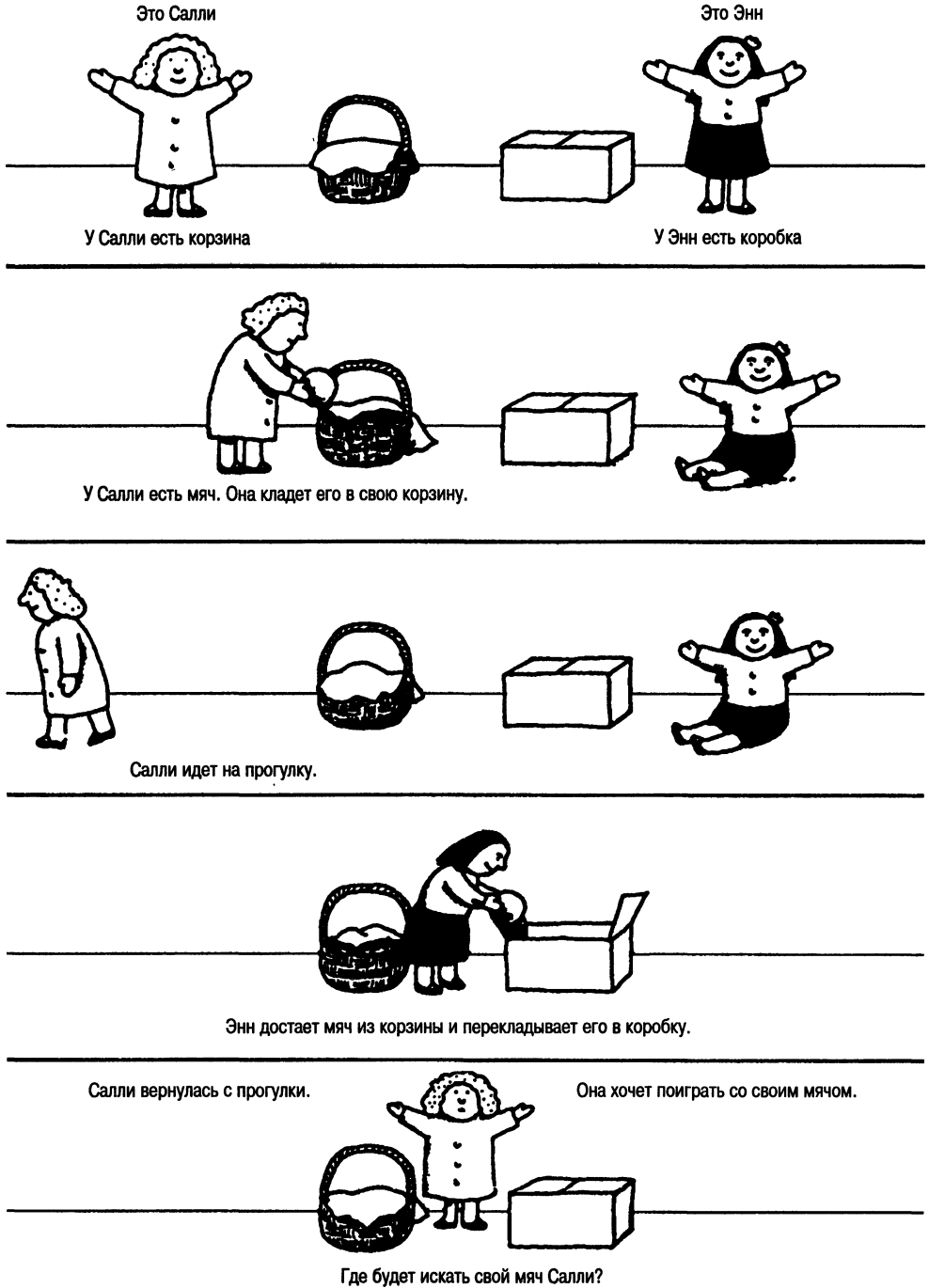


Рис. 11.3. Пример задания «Заблуждение». Чтобы дать правильный ответ, ребенок должен осознать, что мнение — это мысленная репрезентация, которая может не соответствовать действительности. (Источник: *Autism: Explaining the Enigma* (p. 160), U. Frith, 1989, Oxford: Basil Blackwell)

стимулировало разработку новых методов изучения представлений ребенка о психических феноменах, и именно на методологическом аспекте я и сосредоточусь.

Один из подобных методов иллюстрирует рисунок 11.3. Понятие, представление о котором оценивается в задании «Салли/Энн», — **заблуждение** — мнение, не соответствующее действительности. Для того чтобы правильно ответить на вопрос, очевидно, необходимо обладать представлением о том, что люди могут заблуждаться. Чтобы понять, что Салли была бы уверена в чем-то, что не соответствует реальности, что она заблуждалась бы, ребенок должен отвлечься от того, что ему известно о местонахождении мяча. Как уже отмечалось, большинство 3-летних детей не справляется с этой задачей, тогда как к 4–5-летнему возрасту большинство детей вполне на это способно. Задание «Заблуждение» представляет интерес, поскольку позволяет узнать нечто очень важное о понимании ребенком мысленных репрезентаций. Чтобы осознать возможное заблуждения, необходимо понимать, что мнение — это просто мысленная репрезентация, а не непосредственное отражение мира и, как репрезентация, оно может соответствовать и не соответствовать действительности. Таким образом, я могу верить в то же, во что и вы, или во что-то другое; и то, во что мы верим, может соответствовать или не соответствовать действительности. Именно этого, по всей видимости, и не понимает маленький ребенок.

Существует два основных метода изучения представлений о заблуждении. Подход, который иллюстрирует рис. 11.3, называется *заданием с неожиданным местонахождением*; эту процедуру разработали Уиммер и Пернер (Wimmer & Perner, 1983). Другой подход — *заданием с неожиданным содержимым*, которое впервые использовали Хогреф, Уиммер и Пернер (Hogrefe, Wimmer, & Perner, 1986). Допустим, мы показываем ребенку коробку от конфет и спрашиваем, что, по его мнению, должно находиться внутри. Ребенок отвечает «конфеты»; тут мы открываем коробку, показывая, что в действительности в коробке лежат карандаши. Затем мы закрываем коробку, берем игрушку Эрни — персонажа из «Улицы Сезам» и задаем следующий вопрос: «Эрни не видел, что находится в коробке. Что он ожидает там увидеть?» Вы и я (и большинство 4- и 5-летних детей) ответили бы «конфеты» — мы понимаем, что Эрни видел коробку лишь снаружи, и поэтому у него сформировались ошибочные представления о ее содержимом. Большинство 3-летних детей отвечают «карандаши». Как и в задании с местонахождением, маленькие дети не в состоянии абстрагироваться от того, что известно им самим о действительном положении вещей. Дети полагают, что все должны придерживаться единственно верного мнения, которого придерживаются они.

Задание о содержимом может использоваться для оценки и других представлений. Можно спросить ребенка не о мнении Эрни (или некой другой третьей стороны), а о его собственном первоначальном убеждении. Теперь вопрос состоит в том, осознает ли ребенок то, что он сам может заблуждаться. Первая часть процедуры остается без изменений: ребенку показывают коробку от конфет, спрашивают его мнение о ее содержимом и раскрывают коробку, демонстрируя то, что в ней находится на самом деле. Однако проверочный вопрос теперь звучит иначе: «Что ты ожидал увидеть в коробке, пока я ее не открыл?» Можно было бы предположить, что этот вопрос не должен вызывать затруднений; в конечном счете, все, что ребен-

ку нужно сделать, это повторить свой ответ, данный несколько секунд назад. Однако 3-летние дети склонны отвечать «карандаши». Понимание собственного заблуждения кажется столь же затруднительным, что и понимание заблуждений другого. Осознание того, что собственное состояние сознания может изменяться, называется **репрезентационным изменением**.

Мы вскоре вернемся к понятию о заблуждении. Однако прежде приведу еще несколько примеров исследований, проводившихся под рубрикой «теория сознания». Другие задания и соответствующие результаты обсуждаются в ряде других источников (Asington, 1993; Gopnik & Meltzoff, 19976; Lewis & Mitchell, 1994).

Различение видимости и реальности — один из элементарных когнитивных навыков — это понимание того, что внешний вид вещи может не соответствовать тому, что она есть на самом деле. Вы понимаете, к примеру, что соломинка, поставленная в стакан с водой, в действительности не сгибается, входя в воду, — она только выглядит согнутой. Аналогично, вы отдаете себе отчет в том, что вырезанная из белой бумаги бабочка, подставленная под красный светофильтр, не становится красной — она просто выглядит красной. Маленькие дети не сразу приходят к пониманию этого факта. Задание со светофильтром — часто используемая методика, которая обычно проводится следующим образом (например, Flavell, Flavell, & Green, 1983). После предварительной словарной подготовки экспериментатор показывает ребенку белую бумажную бабочку в своем естественном виде, а затем накладывает на нее цветное стекло и задает два вопроса: «Когда ты смотришь на бабочку, она *кажется* красной или *кажется* белой?» «А *по-настоящему*, бабочка на самом деле белая, или она на самом деле красная?» Для того чтобы дать правильный ответ, необходимо различать видимость и реальность, а это как раз то, чего маленькие дети не умеют делать; большинство 3-летних детей отвечает, в обоих случаях руководствуясь внешним видом, то есть они заявляют, что бабочка на самом деле красная, а не просто кажется красной. Оценка с точки зрения видимости без учета реальности называется *феноменалистической ошибкой*. В другом варианте задания на различение видимости и реальности дети делают противоположную ошибку, *реалистическую ошибку*, давая оценку с точки зрения реальности и не учитывая при этом видимость. В этом случае часто используют такой стимул как «поролон/камень»: кусок поролона, который выглядит совершенно как серая каменная глыба. Пощупав кусок поролона, и убедившись в том, что это не камень, дети должны ответить на два вопроса: как это выглядит и что это в действительности? Здесь, что примечательно, 3-летние дети склонны заявлять, что объект — это не просто кусок поролона, но и выглядит как кусок поролона. Узнав о реальности, они не могут больше перейти к противоречащей ей видимости.

Несмотря на содержательные различия, способность осознавать заблуждения и разницу между видимостью и реальностью имеет в своей основе общий когнитивный навык. Для решения обеих задач необходимо умение одновременно учитывать две репрезентации одного объекта или события. В случае с «видимостью и реальностью» ребенок должен представлять стимул выглядящим как одно, но в действительности являющимся чем-то другим. В случае с «заблуждениями» ребенок должен представлять и свое верное мнение (например, мяч в коробке), и заблуждение другого человека (например, мяч в корзине). Это фактически лишь два

из множества достижений, основанных на единой способности рассматривать нечто одновременно в двух ипостасях. Большинство форм принятия чужой точки зрения, к примеру, также подразумевают такую двойственность репрезентации. Так, я должен понимать, что картинка, которую я со своей точки зрения вижу в правильном ракурсе, для сидящего напротив выглядит «вверх ногами».

В отличие от исследований понимания заблуждений и разницы между видимостью и реальностью, последний вид исследований, который мы рассмотрим, не имеет общепринятого названия. Однако, как и все описанные выше методы, он направлен на изучение одного из важнейших компонентов представлений ребенка о мире. Многие исследователи пытаются выяснить, что дети думают об *истоках знаний* — то есть, о том, как формируются наши представления о мире (например, O'Neill, Astington, & Flavell, 1992; Pratt & Bryant, 1990). Осознают ли дети, к примеру, что в формировании наших представлений о мире ключевую роль играет восприятие? Понимают ли они, что тот, кто имеет перцептивный доступ к стимулу, получит о нем информацию, а кто не имеет — не получит? Различают ли они модальности восприятия, понимая, что можно узнать (или они сами узнают) посредством зрения, слуха, осязания? Отдают ли они себе отчет в том, что восприятие — не единственный источник знания, что есть еще, к примеру, логическое рассуждение или получение информации через общение? И относят ли они все свои суждения не только к самим себе, но и к другим людям? Как следует из этого беглого рассмотрения (некоторых из) множества неясных моментов, вопросов слишком много, чтобы пытаться кратко изложить результаты исследований. Однако я отмечу один общий вывод, сделанный на основе работ, посвященных представлениям об истоках знаний; вывод, согласующийся с результатами исследований понимания заблуждений и разницы между видимостью и реальностью. Он заключается в том, что младшие дошкольники имеют весьма ограниченные представления о том, откуда мы черпаем информацию об окружающем нас мире, и часто они ошибаются при выполнении заданий, предельно ясных для любого взрослого. Они, к примеру, нередко заявляют, что всегда умели то, чему их только что научил экспериментатор (Taylor, Ebensen, & Bennett, 1994), и уже несколько секунд спустя после того, как они чему-то научились, нередко бывают не в состоянии ответить на вопрос, что явилось источником информации: зрение, осязание или слух (O'Neill & Gopnik, 1991). И вновь мы видим, что ребенку предстоит проделать большой путь в развитии своих представлений о психике.

Спорные вопросы исследований

Главный вопрос, который теперь перед нами встает, тот же, что и вопрос, к рассмотрению которого мы обращались в этой главе в связи с подходом Пиаже: заложив основы области исследования, очертив ее границы, какие еще исследования могли бы мы провести? Интересно было бы получить ответы на множество вопросов, однако я ограничусь анализом двух, о которых шла речь при обсуждении работ Пиаже — проблемой *оценки* и проблемой *закономерностей*. Оба вопроса мы рассмотрим в их отношении к заданиям о заблуждениях, однако замечания, которые будут сделаны, носят общий характер и относятся не только к исследованиям понимания заблуждений.

В определенном смысле исследование понимания заблуждений интересно по той же причине, что и многие исследования Пиаже: оно позволяет получить удивительную, неожиданную информацию об ограниченности когнитивных возможностей маленьких детей. Так же, как примечательно непонимание младенцами постоянства объектов и непонимание дошкольниками принципов сохранения, удивительно непонимание 3-летнего ребенка того, что человек может заблуждаться. Как и в случае с исследованиями Пиаже, вполне естественная и адекватная реакция на подобное совершенно неожиданное заявление — усомниться в точности оценки. Действительно ли стандартные задания позволяют точно оценить знания маленьких детей о том, каким может быть мнение человека? Вполне вероятно, что это не так, и по тем же причинам, по которым подвергаются критике методы Пиаже: задания сложны в языковом отношении и могут сбить с толку маленького ребенка, контекст проведения обследования непривычен и, мягко говоря, несколько неестествен, а процедуры не вызывают у ребенка интереса и мотивации, необходимых для оптимального выполнения заданий. Чендлер и Хала (Chandler & Hala, 1994, p. 412) подводят итог этим недостаткам следующим образом: «В этих заданиях речь обычно идет о вещах статичных, гипотетических, имеющих отношение к некоему третьему лицу, не актуальных для маленьких испытуемых и не представляющих для них личного интереса». Возможно, 3-летние дети оказались бы более компетентными, если процедуру оценки можно было бы сделать в большей мере ориентированной на ребенка.

Как и в случае с заданиями на сохранение, исследователи разными способами пытаются выявить предпосылки определенных когнитивных навыков, формирующиеся в раннем детстве. Опять-таки используются модификации словесных формулировок заданий. При использовании метода с неожиданным местонахождением, к примеру, вполне вероятно, что маленькие дети истолковывают стандартный вопрос как вопрос о том, где герой *должен* искать, или как вопрос, где герой, *в конце концов*, будет искать; в любом случае, дети ошибаются, указывая на истинное местонахождение. Для исследования этой возможности Сигал и Битти (Siegal & Beattie, 1991) добавили в вопрос потенциально уточняющее слово «сначала» («Где Джейн сначала будет искать своего котенка?»). Другие исследователи изменяют характер ответа, требуемого от ребенка. Было выдвинуто предположение, что маленьким детям предсказать действие, основанное на заблуждении (обычная методика), труднее, чем объяснить основанное на заблуждении действие, которое уже свершилось. Некоторые исследователи (например, Bartch & Wellman, 1989; Moses & Flavell, 1990) в связи с этим сравнивали стандартную процедуру с предсказанием и процедуру с объяснением, когда ребенку предлагалось истолковать ошибочные действия персонажа (например, поиски котенка не в том месте, где он находился). Клементс и Пернер (Clements & Perner, 1994) еще более снижают требования к ответу ребенка. Они следили за траекторией взгляда детей, выполнявших задание с неожиданным местонахождением; их цель состояла в том, чтобы проверить, не смотрят ли дети сначала на то место, где предмет находился до перемещения, даже если они неспособны ответить на стандартный вопрос в вербальной форме (как оказалось, многие дети именно так и делают).

Наверное, чаще всего используется модификация процедуры оценки представлений о заблуждении, включающая в себя элементы *обмана*. Для этого есть и теоретические, и методологические основания. Теоретически понимание обмана предполагает понимание заблуждения, поскольку функция обмана — вводить в заблуждение. Методологически оценка, основанная на обмане, может строиться на естественном опыте розыгрышей и игр, что добавляет непринужденность и личную заинтересованность, которых не хватает стандартным методикам. Исследователи пытаются выяснить представления ребенка об обмане разными способами. Одни включают обман в стандартную процедуру, прося ребенка помочь кого-то «разыграть»; ребенок здесь должен заменить содержимое коробки конфет или переложить игрушку в другое место (например, Sullivan & Winner, 1993). Другие — разрабатывают игры, в которых ребенок, чтобы получить приз, должен обмануть соперника, например неправильно показав место, где лежит вожаемая шоколадная конфета (например, Russell, Mauthner, Sharpe, & Tidswell, 1991). Здесь, очевидно, мотивация следить за ходом мыслей другого человека должна быть максимальной.

Какие выводы были сделаны в исследованиях с модифицированными процедурами оценки? На сегодняшний день это довольно спорный вопрос: разные специалисты по-разному трактуют полученные данные (сравните, Chandler & Hala, 1994; Sodian, 1994). Здесь я ограничусь скромной интерпретацией. По всей видимости, даже маленькие дети имеют некоторые представления о мнении и состоянии сознания другого человека — представления, которые не всегда выявляются в стандартных заданиях. Кроме того, по всей видимости, эти представления на начальных этапах весьма ограничены и отрывочны; в дошкольном возрасте происходит их постепенное совершенствование. Непонимание заблуждений, как и непонимание принципов сохранения, — когнитивный феномен, который действительно имеет место.

Вопрос о закономерностях сходен с вопросом, который обсуждался при рассмотрении теории Пиаже: как соединяются друг с другом различные когнитивные навыки в процессе развития ребенка? Как и в случае с исследованиями Пиаже, этот вопрос представляет интерес и с точки зрения теории, и с точки зрения описания, поскольку в различных концепциях «теории сознания» выдвигаются разные предположения о том, какие следует искать взаимосвязи. Обычно на него пытаются ответить, используя корреляционный подход: предъявляют несколько методик одной и той же выборке детей и оценивают меру согласованности или несогласованности ответов в разных заданиях. В одних исследованиях оценивается согласованность ответов на разного рода задачи, связанные с заблуждениями, к примеру, на задания с неожиданным местонахождением и неожиданным содержимым или на вопросы о себе и о другом человеке в заданиях с неожиданным содержимым (например, Holmes, Black, & Miller, 1996). В других исследованиях анализируется корреляционная связь между результатами выполнения задачи на заблуждение и результатами выполнения других задач о «теории сознания», к примеру, заданий на понимание разницы между видимостью и реальностью, заданий на умение принять чужую точку зрения (например, Gopnik & Astington, 1988). Общий вывод, который можно сделать на основе всех этих исследований, также довольно скромный: разные когнитивные навыки, имеющие отношение к «теории сознания», обычно кор-

релируют между собой; однако эта корреляция далека от абсолютной. Во многих исследованиях корреляции остаются даже при исключении фактора возраста — важное свидетельство того, что они обусловлены не просто возрастными изменениями. Доказательство дивергентной валидности — то есть получение данных о том, что результаты выполнения заданий о «теории сознания» коррелируют между собой в большей степени, чем с результатами выполнения других тестов, — задача будущих исследований.

В перспективе, вопрос о закономерности можно изучать лонгитюдным методом. Лонгитюдное исследование позволяет оценить межвременную связь когнитивных навыков, формирующихся на разных этапах развития. Прогностичен ли некий опыт или некое достижение в 2-летнем возрасте в отношении имеющих с ними общую теоретическую основу когнитивных достижений в 4-летнем возрасте? Сегодня проведение лонгитюдных исследований только начинается (Astington & Jenkins, 1995; Dunn, 1995), однако они, несомненно, окажут существенное влияние на развитие такой области исследования, как «теория сознания».

Резюме

В этой главе рассматриваются четыре подхода к изучению когнитивного развития детей. Работы Пиаже долгое время занимали господствующее положение в области исследования мышления детей, поэтому эту главу открывает обсуждение именно его исследований. Как и исследования младенцев, исследования последующих этапов развития, проведенные Пиаже, отличает стремление выявить элементарные формы представлений посредством использования гибкого «клинического метода» тестирования. Из всех этих форм представлений *сохранение* — осознание того, что при изменении внешнего вида объекта, его количественные характеристики остаются инвариантными, — оказалось особенно интересным и Пиаже, и другим исследователям. Помимо сохранения описываются еще четыре понятия, изучавшиеся Пиаже: *сложение классов*, *транзитивность*, *формально-операциональное мышление* и *принятие чужой точки зрения*.

Работы Пиаже стимулировали дальнейшие исследования, направленные как на опровержение, так и на развитие положений его теории. Обсуждается три основных спорных вопроса в исследованиях когнитивного развития. Первый — вопрос *оценки*: насколько точно с помощью используемой процедуры можно оценить способности ребенка? Пиаже нередко критикуют за недооценку способностей детей, преимущественно по причине излишнего внимания к вербальным заданиям. То, что эта критика небезосновательна, подтверждается результатами исследований двух типов: исследований, в которых использовалось языковое упрощение стандартных заданий Пиаже, и исследований, в которых ситуации оценки придается большая естественность. Второй из спорных вопросов — *закономерности*: как соединяются различные когнитивные навыки? Хотя Пиаже в своей теории говорит о последовательности и синхронии в развитии отдельных навыков, использовавшиеся им экспериментальные процедуры не позволяют получить убедительных доказательств в отношении подобных взаимосвязей. Обсуждаются как общие трудности изучения закономерностей, так и частные недостатки подхода Пиаже. По-

следний спорный вопрос — *когнитивные изменения*: как элементами когнитивной системы становятся новые навыки? Поскольку тема когнитивных изменений не получила должного развития в исследованиях самого Пиаже, описывается один из подходов, разработанных сравнительно недавно: использование *обучающего исследования* для выявления механизмов когнитивных изменений.

Тестовый, или *IQ*-подход к изучению интеллекта детей во многом отличается от подхода Пиаже. Обсуждение этих различий служит в качестве введения к описанию содержания двух ведущих тестов интеллекта: теста Стэнфорд—Бине и *WISC*. Эти тесты представляют собой высокостандартизированные методики, направленные на выявление индивидуальных различий детей. По содержанию они ориентированы на академически-вербальные навыки, важные при обучении в школе, что неудивительно, учитывая прагматические цели тестирования интеллекта. В действительности, именно прогностические возможности этих тестов, включая и корреляционную связь с уровнем школьной успеваемости, всегда служили главным доказательством их валидности.

За описанием *IQ*-тестов следует рассмотрение некоторых теоретических вопросов, представляющих интерес при изучении интеллекта. Первый — вопрос *стабильности*: остается ли неизменным *IQ* ребенка в процессе развития, или же уровень интеллекта может повышаться или снижаться? Для ответа на этот вопрос требуется проведение лонгитюдных исследований; поэтому в этом разделе главы напоминаются некоторые замечания, которые были сделаны в главе 3 при рассмотрении лонгитюдного плана исследования. Еще один существенный вопрос касается *детерминант различий* в *IQ* — в частности, вопрос о том, определяются эти различия генетическими или средовыми факторами. Анализируется два подхода к изучению этого вопроса. Близнецовые исследования строятся на том, что у однояйцевых близнецов уровень генетической общности выше, чем у двуяйцевых; поэтому большее сходство *IQ* у однояйцевых пар рассматривается в качестве свидетельства значимости генетических факторов. В этом отношении особый интерес представляют исследования близнецов, воспитывающихся порознь. Исследования приемных детей дают возможность разделить генетические и средовые факторы, которые смешиваются при нормальной связи между ребенком и родителем. *IQ* приемного ребенка можно сравнить с *IQ* его биологических родителей (с которыми он связан генетически) и *IQ* приемных родителей (с которыми он связан через фактор среды). Как бы ни были значимы близнецовые исследования и исследования приемных детей, они имеют некоторые недостатки, и эти недостатки, наряду с полученными результатами, обсуждаются в представленной главе.

Последние разделы главы посвящены обсуждению двух важнейших направлений в изучении когнитивной сферы детей. Память — это базовый когнитивный процесс, данный с рождения. Ведущим мнемическим процессом у младенцев является, по всей видимости, *узнавание*, которое обычно изучается методом габитуации/дегабитуации или методом предпочтений Фанца. В исследованиях детей старшей возрастной группы основное внимание уделяется процессу *воспроизведения* и попытке объяснить совершенствование памяти, происходящее в период детства. Обсуждаются два подхода к этому вопросу. Исследования стратегий направлены на выявление и оценку эффективности разнообразных *мнемических стратегий*,

которые можно использовать для облегчения работы памяти. В одних случаях использование стратегий наблюдаемо в естественном поведении; в других случаях возникает необходимость либо судить об использовании стратегии по успешности выполнения тестов памяти, либо вызывать использование стратегий экспериментальным путем. В исследованиях *конструктивной памяти* анализируется влияние на мнемические процессы системы общих представлений и развития этой системы. Приводится несколько примеров исследований конструктивной памяти, включая исследования памяти на предложения и памяти на связанные рассказы.

Под термином *теория сознания* подразумевают представления детей о психическом мире. Особого внимания заслуживают представления о *заблуждении*: осознание того, что люди могут придерживаться ошибочного мнения. Чтобы это осознать, необходимо обладать способностью одновременно учитывать две репрезентации, способностью, лежащей в основе еще одного важнейшего навыка: *различения видимости и реальности*. Описывается несколько конкретных методов исследования не только представлений о заблуждении и различения видимости и реальности, но и третьего предмета изучения исследований «теории сознания»: понимания детьми *истоков знаний*. Как и в случае с исследованиями Пиаже, в исследованиях «теории психики» важен вопрос *оценки*; в связи с этим дается описание нескольких альтернативных подходов к изучению представлений о заблуждении. Имеет значение и вопрос о *закономерностях* развития разнообразных навыков, который изучается посредством анализа корреляций между различными навыками, связанными с «теорией сознания».

Упражнения

Одна из привлекательных особенностей заданий Пиаже и заданий, связанных с «теорией сознания», состоит в том, что процедура их использования проста и поэтому очень удобна для осуществления показательных обследований. Это упражнение предполагает, что у вас есть возможность протестировать, по крайней мере, одного ребенка (предпочтительнее — нескольких детей). Если возраст детей от 3 до 5 лет, обратите внимание на методики, имеющие отношение к «теории сознания»; если дети старше — используйте задания Пиаже. В любом случае, прочитайте описание, данное в тексте, а также соответствующие первоисточники, попробуйте применить некоторые методики и сравните полученные вами результаты с результатами, приведенными в литературе. Внимание: прежде чем проводить тестирование, получите осведомленное согласие детей и их родителей.

Исследования, направленные на оценку теории Пиаже в ее части, посвященной зрелому возрасту, свидетельствуют о том, что многие взрослые не всегда успешно справляются с заданиями, направленными на оценку формально-операционного мышления. Достаньте экземпляр книги Инельдера и Пиаже, на которую есть ссылка в тексте, выпишите оттуда как можно больше заданий и предъявите эту батарею тестов выборке ваших сверстников. Как вы считаете, действительно ли подобные задания измеряют способность к гипотетико-дедуктивному рассуждению? Если нет, то как бы вы видоизменили эти задания или что бы вы к ним добавили?

Одним из поводов для критики *IQ*-тестов является то, что с их помощью измеряют только некоторые когнитивные навыки, игнорируя тем самым другие формы когнитивной компетентности. Чаще всего среди аспектов интеллекта, которые игнорируются в этих тестах, называют сферу социального или межличностного интеллекта. Представьте, что вам поручили разработать тест социального интеллекта для младших школьников. Какие бы задания и вопросы вы включили в свою методику? Как бы вы провели ее валидизацию?

Восстановите в памяти свою мнемоническую деятельность за последнюю неделю. В какой мере она отражает процессы, описанные в тексте? Предположим, некий исследователь захотел изучить вашу ежедневную мнемоническую деятельность. Каким образом он мог бы осуществлять сбор данных?

Социальное развитие

Как и глава 11, эта глава посвящена развитию в постмладенческом возрасте. В центре внимания теперь — социальное развитие ребенка, тема не менее глобальная и не менее сложная с методологической точки зрения, чем тема когнитивного развития. Изложение материала вновь будет избирательным; наша цель состоит в том, чтобы рассмотреть несколько важнейших аспектов более или менее подробно, а не множество аспектов — поверхностно.

Глава делится на две большие части. Первая часть посвящена главным аспектам социального развития. Два из них, привязанность и темперамент, рассматривались в главе 10. В этой главе мы добавим к ним еще три: развитие представлений о нравственных нормах и развитие нравственного поведения; развитие Я-концепции, а также развитие половой типизации и половых различий.

В отношении любой исследовательской проблематики должны быть разрешены два методологических вопроса. Первый — как измерять интересные нам аспекты развития ребенка — разнообразные формы нравственного поведения, проявления половой типизации и т. д. Рассмотрению этого вопроса посвящена первая часть главы. Сюда включено обсуждение практического применения принципов, изложенных ранее (например, разница между лабораторным исследованием и исследованием в полевых условиях), а также конкретных методик, которые используются в разных областях исследования.

Второй большой вопрос касается детерминант — почему все происходит именно так? Почему в социальный репертуар поведения ребенка входят определенные формы поведения, и чем объясняются возрастные и индивидуальные различия в поведении? Именно этому вопросу посвящена вторая часть данной главы. Как и прежде, представленный материал объединяет в себе общие и более специфические вопросы.

Нравственное развитие

«Нравственное развитие» — глобальная тема, охватывающая множество аспектов развития ребенка. Фактически, в том виде, в котором нравственность рассматривают специалисты в области психологии развития, это действительно глобальная

тема, поскольку под этой рубрикой изучаются различные аспекты развития. Поэтому первый раздел — самый большой из разделов, составляющих эту главу.

Традиционно выделяют три компонента нравственности: поведенческий, эмоциональный и когнитивный. Излагая материал, я также придерживаюсь этого традиционного деления, отводя для каждого из компонентов отдельный подраздел. В каждом из этих подразделов основное внимание уделяется первому из указанных выше общих вопросов: как измерять интересующие нас аспекты развития? Второй из общих вопросов, как исследовать детерминанты этих аспектов, анализируется в заключительной части данной главы.

Поведенческий аспект

Поведенческий аспект нравственности связан с проявлением нравственности в поведении, с поведением, которое нравственно, одобряемо и социально приемлемо. Каковы именно критерии для определения того, «нравственно» ли поведение, — сложный и весьма спорный вопрос, обсуждение которого не входит в нашу задачу. Отмечу, однако, что одним из компонентов практически любого определения является указание на то, что нравственное поведение должно быть, по крайней мере, частично, обусловлено внутренним побуждением, а не только непосредственным внешним давлением. Ребенок, который не списывает, потому что над ним стоит учитель, не ведет себя нравственно (или безнравственно, если уж на то пошло), как и ребенок, который делится конфетами с младшим братом, потому что мать грозит его отшлепать, если он этого не сделает. С другой стороны, поведение ребенка, которому предоставляется некоторая степень свободы и который решает не списывать или решает поделиться, можно назвать нравственным.

Еще один характерный критерий при оценке нравственности поведения — бескорыстность. Поведение человека можно однозначно назвать нравственным, если оно противоречит личным интересам (или по крайней мере не корыстно). Так, ребенок, который делится конфетами, потому что не любит этот сорт конфет, не обязательно ведет себя нравственно, равно как и ребенок, который не списывает, потому что в этом нет необходимости. С другой стороны, ребенок, который дарит последнюю конфету, которую он с большим удовольствием съел бы сам, демонстрирует действительно нравственное поведение.

Заметьте, что эти моменты, касающиеся определения, имеют непосредственное методологическое значение. Мы должны изучать ситуации, либо естественные, либо экспериментально созданные, в которых у ребенка появляется возможность повести себя нравственно, не руководствуясь своекорыстными интересами и не по принуждению. По мере изложения я буду приводить примеры подобных ситуаций.

Какие конкретно формы поведения попадают в категорию поведенческого аспекта нравственности? Существует огромное множество таких форм, которые можно приблизительно разделить на два больших класса, соответствующих двум приведенным выше примерам. В одних случаях вопрос состоит в том, может ли ребенок не поддаваться искушению совершить проступок. Это, к примеру, случай со списыванием, — часто используемая методика оценки нравственного поведения у детей школьного возраста. Списывание — это, фактически, частный случай такого конструкта, как устойчивость перед соблазном. В тестах устойчивости перед соблазном создается конфликт между склонностью действовать в личных интересах (например, списывать с целью получения более высокой оценки) и некими санк-

циями или запретами, наложенными на это поведение (например, правилами, запрещающими списывать, или неодобрением учителя, родителя). Важно, устоит или не устоит ребенок перед соблазном. Очень часто для измерения устойчивости маленьких детей используется тест «Запрещенные игрушки», который мы вскоре опишем.

Долгие годы изучение нравственного поведения было ориентировано главным образом на ситуации, когда ребенку приходилось противостоять искушению. В последнее время внимание к негативным ситуациям дополнил возрастающий интерес к позитивной стороне нравственности: сможет ли ребенок не только устоять перед соблазном совершить проступок, но и повести себя благородно? Активное социально одобряемое поведение называется **просоциальным**. Пример просоциального поведения — проявление щедрости. Просоциальным поведением можно также назвать сочувствие огорченному человеку или предложение помощи для смягчения его горя. Более формально просоциальное поведение определяется как «действия, цель которых — помочь или принести пользу какому-либо человеку или группе людей, при том что совершающий эти действия не рассчитывает на награду» (Mussen & Eisenberg, 1977, pp. 3–4). Еще раз обратите внимание на отсутствие внешнего контроля нравственного поведения.

Какую бы форму поведения мы ни решили изучать, для сбора данных можно использовать несколько подходов. В общих чертах эти подходы были описаны в главе 6; здесь же мы рассмотрим их конкретное применение к исследованию нравственного развития.

Один из путей — найти некие естественные «полевые» условия и провести натуралистическое наблюдение интересующего нас поведения. Примером может служить исследование Барретт и Йарроу (Barrett & Yarrow, 1977). Они изучали 5–8-летних детей, посещавших летний детский лагерь. Использовался метод обсервативных временных срезов, при этом каждого ребенка в целом наблюдали в течение восьми 15-минутных периодов; каждый период наблюдения, в свою очередь, был разделен на три 5-минутных блока. При наблюдении фиксировались в том числе и случаи просоциального поведения, определявшегося как:

«Попытки удовлетворить потребность другого человека в физической или эмоциональной поддержке. К ним относятся акты утешения (физически или вербально выражаемого сочувствия или подбадривания), дарения (предоставления возможности использовать предметы или рабочее пространство) и помощи (физического содействия или предложения физического содействия)» (p. 476).

Кроме того, фиксировались ситуации, когда имелся шанс повести себя просоциально, — то есть сигналы со стороны других детей, свидетельствующие о том, что они нуждаются в утешении, помощи или в том, чтобы с ними поделились.

Очевидное преимущество подходов, аналогичных подходу Барретт и Йарроу, отмечалось в главе 6. При натуралистическом наблюдении мы непосредственно видим то, что пытаемся объяснить: естественное поведение в естественных условиях. Еще одно преимущество, касающееся конкретно этого исследования, — оценка не только случаев самого просоциального поведения, но и шансов его продемонстрировать. Типичным недостатком натуралистического наблюдения является то, что ситуации, провоцирующие определенное поведение, неодинаковы у разных детей. Возможно, к примеру, что у ребенка низкие показатели по шкале «утешение» лишь потому, что на момент наблюдения у него не было возможности никого

утешить. Эта проблема является следствием недостатка контроля в натуралистическом исследовании: мы не можем создать одинаковые условия для всех детей, а должны принимать обстоятельства такими, какими они сложились. Чтобы разрешить эту проблему, можно учитывать подобные обстоятельства у каждого ребенка или соответствующим образом модифицировать процедуру анализа, что и сделали Барретт и Йарроу, приняв во внимание возможности для демонстрирования просоциального поведения. Еще один способ, который также использовали Барретт и Йарроу, — сбор данных о специфических моделях поведения каждого ребенка. Логика при этом аналогична рассматривавшейся ранее логике отбора испытуемых: чем больше выборка или чем больше данных о специфических моделях поведения каждого ребенка, тем выше вероятность компенсирования случайных вариаций и, следовательно, получения репрезентативных результатов.

Как мы указывали выше, натуралистический подход обладает не только достоинствами, но и недостатками. Затруднительно обеспечение точности измерений, особенно в тех случаях, когда наблюдение необходимо проводить в неконтролируемых полевых условиях. В исследовании Барретт и Йарроу надежность, основанная на общности оценок наблюдателей, в действительности была довольно умеренной, в пределах 0,7–0,8. Еще одна проблема — возможность влияния наблюдателя на наблюдаемое поведение. Да, в ситуации свободной игры, как в исследовании Барретт и Йарроу, присутствие взрослых не является чем-то необычным, и к нему легко привыкнуть. Тем не менее в этом отношении не следует быть слишком оптимистичными. Осознание своей роли наблюдаемого, в действительности, может особенно сильно повлиять на поведение, когда речь идет о нравственности, о сфере, которая диктует определенные нормы того, как «следует» себя вести в присутствии взрослого (то есть следует делиться, следует помогать и т. д.). Не исключено, что вероятность просоциальных действий ниже, когда поблизости нет взрослого.

Рассмотрим еще один пример натуралистического обсервативного исследования. Это необычно смелое исследование, в котором была предпринята попытка минимизировать описанные только что трудности. Зан-Вакслер, Радке-Йарроу, Вагнер и Чэпмен (Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992) изучали реакцию на огорчение другого человека у 1–2-летних детей — само по себе довольно смелое предприятие, поскольку дети этого возраста представляют собой трудную и недостаточно изученную группу. Исследователей особенно интересовали просоциальные реакции, вызванные огорченностью другого человека, такие как выражения сочувствия или предложения помощи. Эти реакции изучались в лонгитуде в течение одного года, что, учитывая трудности лонгитюдного исследования, является еще одним свидетельством смелости авторов. Наконец, наблюдение вели не обычные ассистенты исследователя, а матери детей. Перед началом исследования матери проходили курс интенсивной подготовки, обучаясь производить наблюдение и фиксировать результаты так, как это требовалось исследователям. Обсервативный метод представлял собой сочетание выборочного анализа поведенческих событий и нарративной записи: всякий раз, когда кому-то случалось огорчаться, мать надиктовывала на магнитофон описание случившегося, предшествовавших событий и реакции ребенка на чужое горе. Использование этого подхода дало возможность выявить ранние формы сочувствия и просоциального поведения.

ния в возрастной группе, в которой, как считалось, такие реакции невозможны (см. также Radke-Yarrow & Zahn-Waxler, 1984; Zahn-Waxler & Radke-Yarrow, 1982).

Наверное, нет необходимости говорить о том, что при проведении таких исследований возникает много трудностей, как практических, связанных с обеспечением сотрудничества, так и более существенных, связанных с обеспечением точности отчетов матерей. Однако этот подход имеет и сильные стороны, особенно если сравнить его с обычными обсервативными исследованиями. Стоит упомянуть два преимущества. Во-первых, сводится к минимуму влияние присутствия наблюдателя, поскольку мать является естественной частью среды, в которой воспитывается ребенок. Хотя наблюдение все же ограничивается случаями, свидетельницей которых становится мать, проблема реактивности в ее обычном виде не стоит. Заметьте, что это пример «включенного наблюдения», о котором говорилось в главе 4. Во-вторых, возможности для наблюдения значительно шире, а объем информации значительно больше, чем в типичном обсервативном исследовании. Поэтому мы можем быть гораздо увереннее в том, что получили репрезентативные данные о поведении ребенка.

Второй подход, который мы рассмотрим, — лабораторное выявление определенного вида поведения. В этом случае наблюдение производится в структурированной обстановке лаборатории, в обстановке, создающей возможности для проявления и измерения конкретных форм нравственного поведения, интересующего исследователя. Проанализируем два наглядных примера этого подхода: первый — ориентированный на негативный аспект нравственности, связанный с сопротивлением соблазну, а второй — ориентированный на позитивный, просоциальный аспект.

Как отмечалось ранее, оценка устойчивости ребенка к соблазну обычно производится с помощью теста «Запрещенные игрушки». Типичный пример его использования — одно из первых исследования Парке (Parke, 1967). В исследовании Парке испытуемых (1- и 2-классников) по одному приглашали в передвижную лабораторию, помещенную в трейлере, и предлагали им порисовать в течение 10 минут. По окончании фазы рисования ребенка усаживали за стол, на котором стояло пять игрушек, и давали следующую инструкцию:

Можешь посидеть здесь. [Э. откидывает накрывающую игрушки ткань]. Эти игрушки предназначены для другого человека, поэтому ты их не трогай. Если будешь хорошим мальчиком (девочкой) и не будешь трогать игрушки, мы с тобой потом поиграем. Я кое-что забыл в школе и должен сходить и принести это. Пока меня нет, можешь полистать эту книгу. Я закрою дверь, так что никто тебя не побеспокоит. Когда я вернусь, то поступчу, чтобы ты знал(а), что это я (p. 1105).

Ребенка на 15 минут оставляли наедине с игрушками и следили за его поведением через одностороннее зеркало. Вопрос, конечно, состоял в том, устоит ли ребенок перед соблазном поиграть с игрушками, и если нет, как скоро он нарушит уговор и как далеко зайдет в его нарушении. Оказалось, что дети в разной степени способны не поддаваться искушению и что эта способность определяется рядом переменных. Одну из этих переменных мы рассмотрим позже, при обсуждении исследований нравственных рассуждений.

Второй пример лабораторного исследования имеет отношение к просоциальной способности делиться (см. Barnett, King, & Howard, 1979). Это исследование иллюстрирует один из наиболее популярных методов изучения этой способности. Испытуемые, учащиеся начальных классов, в награду за выполнение заданий на память получали 30 призовых фишек. В конце исследования, когда все дети уже прошли тестирование, фишки можно было обменять на призы. Однако прежде чем отпускать ребенка с его фишками, экспериментатор показывал ему копилку и сообщал, что

«Многие дети, такие же как и ты, ходят в школу. Но у них не будет возможности поучаствовать в исследовании. Потом, если захочешь, можешь поделиться с этими детьми, бросив в копилку несколько фишек. Ты не обязан делиться, но если хочешь, можешь это сделать» (Barnett et al., 1979, p. 165).

Затем на некоторое время ребенка оставляли одного, чтобы, предоставленный сам себе, он смог решить, делиться ему или нет. Тот факт, что в копилке уже были фишки, усиливал впечатление анонимности. Однако это впечатление было обманчивым: поведение ребенка контролировалось ничуть не меньше, чем в описанном выше исследовании с запрещенными игрушками. Подсчитав количество фишек после ухода ребенка, экспериментатор имел возможность оценить «вклад» каждого из детей.

Лабораторный подход имеет ряд преимуществ. В сравнении с натуралистическим наблюдением, проведение исследования в лабораторных условиях — намного более эффективный способ сбора информации о поведении. Вместо того чтобы, иногда часами, ждать, когда ребенок поведет себя соответствующим образом, мы создаем ситуацию, актуализирующую интересующее нас поведение (когда ребенок выбирает, поддаваться ему соблазну или нет, делиться или нет и т. д.), в какие-то минуты. Более того, мы можем создавать равноценные условия для всех испытуемых. Как отмечалось, при проведении классического натуралистического исследования неравноценность ситуаций, актуализирующих некое поведение, может создать определенные трудности. И при этом мы можем произвести точное и объективное измерение. К примеру, не представляет труда зафиксировать, играет ли ребенок с игрушками, с которыми ему запретили играть, и как скоро он начал это делать. Еще проще подсчитать, сколько фишек бросил ребенок в копилку.

Стоит отметить еще одно преимущество лабораторного подхода. С момента появления таких методов, как тест «Запрещенные игрушки» и «Пожертвование» началась разработка их многочисленных вариантов. Например, мы можем узнать, что происходит с устойчивостью, изменяя привлекательность игрушек, категоричность запрета или доступность альтернативных занятий. В исследовании способности делиться мы можем изменять количество и привлекательность того, чем приходится делиться, то, насколько знаком или симпатичен человек, с которым делятся, наличие или отсутствие экспериментатора при акте «дарения». Иными словами, мы можем целенаправленно производить экспериментальные манипуляции с огромным количеством потенциально значимых независимых переменных.

Недостатки лабораторного подхода, как указывалось в главе 6, можно объединить одним словом — искусственность. Грусек (Grusec, 1982) подводит итог тому, что обычно вызывает критику этого подхода:

«Есть ли что-то общее между способностью устоять перед соблазном поиграть в условно запрещенные игрушки и способностью не поддаться искушению солгать, нарушить правила игры или украсть? Связано ли вообще жертвование только что выигранными фишками неизвестным обделенным детям с оказанием помощи другу, действительно попавшему в затруднительную ситуацию?» (р. 259)

Сомнения вызывает внешняя валидность — достаточно ли высок уровень сходства лабораторных и жизненных ситуаций, чтобы можно было делать какие-либо обобщающие выводы. В любой лабораторной методике легко обнаружить недостатки и факторы, которые могут негативно отразиться на валидности. Я вернусь к этому вопросу, когда мы будем обсуждать лабораторные исследования детерминант нравственного поведения.

С другой стороны, следует еще раз подчеркнуть момент, о котором говорилось в главе 6. Лабораторные и полевые условия следует рассматривать на континууме, а не как дихотомию; экспериментально смоделированные ситуации не обязательно должны радикально отличаться от естественного опыта ребенка. В действительности, нетрудно доказать, что нравственное поведение, в отличие от большинства представляющих интерес для психологов феноменов, поддается естественной, не приводящей к эффектам реактивности оценке в лабораторных условиях. Например, ребенка в естественной манере просят поделиться — высказанное «между прочим» пожелание, внешне совершенно не связанное с задачами исследования. Аналогично, оказание помощи может оцениваться в ситуации, когда экспериментатор «случайно» роняет коробку с теннисными мячами и смотрит, помогает ли ребенок их собирать. Сочувствие может оцениваться в ситуации, когда экспериментатор якобы прищемил палец дверцей шкафчика и отмечает реакцию ребенка (Yarrow & Waxler, 1976). Если искусно произвести эти эксперименты (большое «если»), они будут вполне естественными и информативными.

Отмечу в завершение, что иногда экспериментальное выявление некоторых форм поведения можно осуществить в естественных условиях. Такие попытки были предприняты, к примеру, в описанном ранее исследовании (Zahn-Waxler et al., 1992). Помимо ситуаций огорчения, складывавшихся естественным образом, для оценки реакции ребенка использовались ситуации притворного огорчения матери, якобы ударившейся ногой. Здесь мы используем сразу два преимущества исследований типа «полевые условия — полевые условия», о которых говорилось в главе 6: экспериментальный контроль, необходимый для обеспечения внутренней валидности, плюс естественные условия как фактор повышения внешней валидности.

Последний метод, который мы рассмотрим, — метод рейтинга (или вербально-го отчета). Здесь вновь исследуется естественное поведение в естественных условиях как противоположность экспериментально выявляемому поведению в обстановке лаборатории. Для осуществления измерений вновь привлекается наблюдатель в противоположность автоматической или полуавтоматической регистрации, которая нередко производится в лаборатории. Однако рейтинг носит значительно более обобщенный, оценочный характер и не имеет такой тесной связи с фактическим поведением, как наблюдение. При наблюдении стоит задача зафиксировать специфику текущего поведения, обычно посредством четкой и объективной сис-

темы регистрации, используемой подготовленными наблюдателями. При использовании метода рейтинга задача состоит в выявлении общих характеристик ребенка — насколько он честен или нечестен, щедр или скуп и т. д. Эти характеристики — не непосредственные поведенческие проявления; о них судят, обычно ретроспективно, используя наблюдения, производимые, как правило, в естественных условиях и несистематически, кем-то, кто хорошо знает ребенка. Этот кто-то — чаще всего учитель или родитель. В исследованиях с детьми старшего возраста это может быть другой ребенок или сам испытуемый. В любом случае наша оценка основана на том, какого мнения о ребенке кто-то, кто его знает.

Рассмотрим несколько примеров. В исследовании Рутерфорд и Массен (Rutherford & Mussen, 1968) рейтинговую оценку щедрости мальчиков дошкольного возраста производили воспитательницы детского сада. Каждой воспитательнице выдавали карточки с именами мальчиков из ее группы и просили разделить эти карточки на пять стопок. В первую должны были войти те из мальчиков, «которых я знаю как самых эгоистичных и наименее щедрых». Три промежуточные категории: для тех, «кто щедр не более остальных, то есть не слишком щедр, но и не ведет себя слишком эгоистично». Сортировка карточек — достаточно распространенный метод получения рейтинговых оценок. Используются карточки или нет, рейтинг подразумевает оценку выраженности у ребенка некоторой характеристики — то есть оценку не только того, присутствует или отсутствует некое качество, но и того, в какой степени ребенок щедр, честен и т. д. Рейтинговая оценка ребенка, как правило, по инструкции или по умолчанию дается на основе сравнения ребенка с другими детьми.

Как отмечают Массен и Айзенберг-Берг (Mussen & Eisenberg-Berg, 1977), рейтинговые оценки воспитателей имеют наибольшее значение при изучении детей дошкольного возраста, поскольку с ними воспитатели общаются часто, при разных обстоятельствах и в неструктурированных ситуациях. В более структурированной обстановке школы общение учителя с детьми недостаточно разносторонне и поэтому не позволяет осуществить рейтинговую оценку разнообразных качеств. Однако к школьному возрасту ценным источником информации может стать мнение сверстников. Мартин Хоффман, в частности, использовал метод «социометрического» рейтинга сверстников при проведении ряда исследований, направленных на оценку различных аспектов нравственного развития (Hoffman, 1975; Hoffman & Saltzstein, 1967). В одном из исследований (Hoffman, 1975) измерялось сочувствие другим: ученикам 5 классов было предложено назвать трех одноклассников одного с ними пола, которые «внимательнее остальных относятся к чувствам других ребят и стараются никого не обижать», а также «вступаются за тех, кого высмеивают или обзывают».

Как и другие методы, рейтинг имеет свои плюсы и свои минусы. Если говорить о недостатках, наиболее очевидный из них заключается в том, что рейтинг не имеет непосредственной связи с фактическим поведением; это интерпретационный отчет о поведении. Есть множество причин, по которым рейтинговые оценки могут создать о ребенке неверное представление: недостаточные возможности для наблюдения, наблюдение лишь в определенных ситуациях, неправильное истолкование инструкций, забывчивость, намеренное искажение фактов, следование

стереотипам. Кроме того, рейтинг, предполагающий выведение общих характеристик, — неудачный способ изучения непосредственных детерминант некой формы поведения (вспомним разделение свойство/состояние, рассматривавшееся в главе 4). Если говорить о преимуществах, рейтинговые оценки, данные человеком, который действительно хорошо знает ребенка, позволяют получить всестороннюю информацию, недоступную при использовании других методов. Лабораторные методы неизбежно ограничены скудными данными о зачастую нетипичном поведении ребенка в короткий промежуток времени. Наблюдение в естественных условиях, как, например, в работе Zahn-Waxler et al., 1992, несмотря на все усилия исследователей, также, как правило, ограничено определенными ситуациями и определенными формами поведения. Рейтинг основан на многоплановых данных, включающих в себя практически все, что знают о поведении ребенка родители, учитель или друг. (Сопоставление метода рейтинга и наблюдения можно найти в работе Cairns & Green, 1979.)

Наиболее общий вывод из сказанного в этом разделе уже был сформулирован в главе 6. Все методы имеют свои недостатки. Поэтому для решения любой исследовательской задачи (в нашем случае, для изучения нравственного развития) необходимо направленное использование как можно большего количества методов.

Эмоциональный аспект

Перейдем теперь к рассмотрению эмоционального аспекта нравственности — к тому, какие чувства испытывает ребенок в ситуациях, значимых с точки зрения нравственности. Наибольший интерес исторически вызывало негативное чувство вины — то есть неприятное, связанное с самоосуждением чувство, переживаемое человеком в ситуации, когда он совершил (или только намеревается совершить) что-либо дурное. В последнее время все больший интерес вызывает такое позитивное чувство, как эмпатия, то есть склонность сопереживать эмоциональным реакциям других людей (к примеру, печалиться, когда кому-то грустно).

Эмоции, по определению, феномен внутреннего мира и, будучи таковыми, представляют для исследователя серьезную методологическую трудность. Как определить, что чувствует ребенок? До недавнего времени за ответом на этот вопрос чаще всего обращались к самому ребенку — то есть пытались получить некий вербальный отчет о чувствах, переживаемых в разных ситуациях. Начнем с этого подхода и рассмотрим его применение к чувству вины, а затем приведем два примера оценки эмпатии. После этого обратимся к некоторым из последних попыток перейти от вербальных отчетов к непосредственному измерению эмоциональных реакций.

Самый распространенный способ изучения чувства вины заключается в предъявлении рассказов, в которых главный герой (который обычно имеет много общего с испытуемым) совершает какой-то проступок. Рассказ обрывается на том месте, где описывается проступок, а ребенка просят досочинить окончание истории. Особый интерес представляют окончания, в которых описываются эмоциональные реакции главного героя, в том числе и переживание чувства вины. Интересным может оказаться и дальнейшее поведение героя: убегает он, прячется, признается в своем проступке, пытается загладить свою вину или действует как-то иначе?

Какая бы реакция ни приписывалась персонажу, предполагается, что она свидетельствует о том, что чувствовал бы и как бы себя вел в подобной ситуации сам ребенок. Поскольку предполагается, что ребенок «проецирует» собственные чувства на персонажа, такого рода методики называются проективными. Примеры двух таких рассказов, предназначенных для оценки чувства вины, приведены во вставке 12.1.

Метод незаконченных рассказов имеет ряд преимуществ. Его довольно легко использовать, обработка результатов также не представляет трудностей, кроме того, этот метод позволяет оценить широкий диапазон реакций в ситуациях разной степени тяжести проступка; такого разнообразия практически невозможно достичь при использовании любого другого подхода. Главное, что вызывает возражения в этом методе, конечно же, допущение, лежащее в его основе: допущение, что ребенок реагирует, проецируя собственные эмоции на героя рассказа (а не в соответствии с ожиданиями взрослого). Исследования с использованием подобных методик не оставляют сомнений в том, что это допущение нередко оправдано; однако насколько оно оправдано в целом — спорный вопрос. Следует отметить, что рассказы, подобные тем, которые представлены во вставке 12.1, можно использовать и более директивно, напрямую спрашивая ребенка, что бы он чувствовал или думал, будь он на месте героя (например, Thompson & Hoffman, 1980). Тем не менее с помощью этой методики можно получить лишь вербальный отчет об эмоциях, а не конкретные показатели.

На вербальных отчетах основаны и многие исследования эмпатии. Одна из методик оценки эмпатии у детей, к примеру, Тест аффективных ситуаций на эмпатию (Affective Situation Test for Empathy) (Feshbach & Roe, 1968). Этот тест включает в себя восемь последовательностей из слайдов, изображающих 6–7-летних детей в различных аффективно окрашенных ситуациях. Два сюжета иллюстрируют радость, два — грусть, два — страх и два — гнев. В «радостном» сюжете, например, показан ребенок (мальчик или девочка, в зависимости от пола испытуемого), празднующий день рождения. Один из «грустных» сюжетов связан с потерей собаки; ниже представлены комментарии к картинкам:

Слайд 1 Вот мальчик и его собака. Этот мальчик везде ходит со своей собакой, но иногда собака пытается убежать

Слайд 2 Вот собака снова убегает

Слайд 3 Теперь мальчику ее не найти; может быть, она пропала навсегда (р. 135)

Процедура, которую использовали Фешбах и Ро, довольно проста. После показа каждого сюжета ребенка спрашивали: «Что ты чувствуешь?» Оценка производилась на основе соответствия чувств ребенка и персонажа. Использовались две системы оценки. Как проявление специфической эмпатии рассматривалось полное совпадение чувств ребенка и персонажа, то есть ребенок должен сообщить, что чувствует радость при просмотре сюжета о дне рождения, грусть — при просмотре сюжета о потерянной собаке и т. д. Как проявление неспецифической эмпатии рас-

Вставка 12.1

Примеры проективных рассказов, используемых для оценки чувства вины

Ранним вечером Боб со своим другом бегут по улице, спеша на главный баскетбольный матч сезона. Боб весь в нетерпении, так ему хочется посмотреть игру. Она начнется через пять минут, и мальчики не хотят пропустить ни одного момента. Там будут все ребята. По пути они замечают маленького мальчика, который растерянно смотрит по сторонам. Он бормочет чье-то имя. Боб и его друг — единственные, кто находится поблизости. Они не знают, что это за ребенок. «Смотри, похоже, этот малыш потерялся, — обращается Боб к своему другу. — Может, походить, поспрашивать, чей он? Это займет всего несколько минут». «Да ну. У нас свои дела. Мы же не можем опоздать на игру. И потом, родители наверняка сами его скоро найдут, и с ним будет все в порядке. Пошли, друг ты мне или не друг?» Наконец Боб говорит: «Ну ладно. Наверное, ты прав. Родители сами его скоро найдут». Они пришли на игру вовремя и с удовольствием посмотрели весь матч. На следующее утро Боб выехал покататься на велосипеде и заметил вывешенную газету. На фотографии был изображен тот самый маленький мальчик; которого они видели вчера. В газете сообщалось, что днем за ребенком присматривала соседка. Пока она была в парикмахерской, 4-летний малыш оставался на улице. Она велела ему поиграть и дожидаться ее. Но ребенок пошел гулять и потерялся. Пока соседка его искала, мальчик, перебегая улицу, попал под машину. В газете сообщалось, что по дороге в больницу мальчик скончался.

Арт и его друзья отдыхали на школьном пикнике. На пикнике проводилось много разных соревнований. Арт любил принимать участие в соревнованиях и ужасно хотел победить. Его друзья выигрывали соревнования по прыжкам, поиску кладов и бегу. Арт был одним из тех ребят, кому так и не удалось ничего выиграть. В нескольких конкурсах он даже оказался на последнем месте. «Может быть, я выиграю в плавании. Это главное соревнование!» — размышлял он. Когда пришло время заплыва, Арт стал думать, как бы ему выиграть. Соревнующиеся должны были доплыть под водой до большого белого буйка и обратно. Вся дистанция составляла около 25 футов. Арт знал, что никто не заметит, если он, не доплыв до буйка, повернет обратно, потому что в пруду купалось много людей и вода стала немного мутной. Итак, Арт проплыл лишь часть дистанции, повернул обратно и пришел первым. Все стали поздравлять его с победой. Когда к финишу пришли другие пловцы, они стали хвалить Арта. Никто не заметил, как он повернул. Ему подарили ленту, и никто не подозревал, что он далеко не самый лучший пловец.

Источник: «*Conscience, Personality, and Socialization Techniques*», M.L.Hoffman, 1970, *Human Development*, 13, p. 98.

смаatrивалось совпадение эмоций ребенка и персонажа лишь по знаку. Так, в случае с потерянной собакой в зачет идут любые негативные переживания, например, «плохо», «испугался».

Вслед за методикой для взрослых (Mehrabian & Epstein, 1972) появилась методика оценки эмпатии у детей (Bryant, 1982). Процедура Брайант проще в проведении и обработке результатов, чем процедура Фешбах и Ро (Feshbach & Roe, 1968), и предназначена для более широкого возрастного диапазона испытуемых. Методика состоит из 22 вопросов, представленных в табл. 12.1. Вопросы описывают позитивные и негативные эмпатические реакции на разнообразные ситуации. От маленьких детей требуется лишь ответ «да/нет». Со старшими детьми и подростками может использоваться 9-балльная рейтинговая шкала для более точного указания на степень согласия или несогласия с определенным утверждением.

Хотя методики вышеупомянутых авторов имеют определенную ценность, они страдают тем же недостатком, что и упомянутый ранее метод неоконченных рассказов для оценки чувства вины: они измеряют вербальные отчеты об эмоциях, а

не сами эмоции. Трудность, конечно же, в том, чтобы разработать методику непосредственной оценки эмоционального состояния. В последние годы в некоторых исследованиях изучается два альтернативных индекса эмоций: выражение лица и физиологические изменения. Примером может послужить исследование, которое провели Айзенберг и др. (Eisenberg et al., 1988). В этом исследовании дошкольникам и 2-классникам показывали три фильма, в которых дети были изображены в эмоционально окрашенных ситуациях: мальчик и девочка, застигнутые грозой, маленькая девочка, у которой только что умерла ее птичка, и девочка-инвалид, пытающаяся встать на ноги и самостоятельно пойти. Во время просмотра у детей контролировались частота сердечных сокращений и выражение лица. После просмотра каждого из фильмов детей, кроме того, просили рассказать, какие чувства они испытывали. Ожидалось, что эмпатической реакцией на переживания персонажей первого фильма будет страх или тревога; а второй и третий фильм вызовут грусть и сожаление. Именно это и было обнаружено, хотя у старших испытуемых реакции носили более выраженный характер. Большинство детей сообщали, что они испытывали страх при просмотре первого фильма; а реакция грусти, напротив, была более характерна при просмотре фильма об умершей птице. Анализ выражения лица дал ту же картину: выражение, свидетельствующее о переживании страха при просмотре фильма о грозе, и выражение, свидетельствующее о грусти и сожалении при просмотре двух других отрывков. Наконец, изменения частоты сердечных сокращений также зависели от содержания фильма: частота сердечных сокращений повышалась при просмотре первого фильма и снижалась при просмотре двух других. Снижение частоты сердечных сокращений, как отмечалось в главе 10, — один из признаков состояния внимания; авторы исследования заявляют, что этот аспект реакции на грусть другого человека — свидетельство ориентированного вовне, искренне сочувственного переживания. В подтверждение авторы ссылаются на данные других своих исследований (Eisenberg et al., 1989), указывающих на то, что снижение частоты сердечных сокращений при встрече с кем-либо, нуждающимся в поддержке, связано с повышенной готовностью оказать помощь.

Следует подчеркнуть, что подход Айзенберг к измерению эмоциональных реакций воплощает в себе идею о необходимости использования конвергентных операций, о которой я периодически напоминаю в этой книге. Он, кроме того, служит примером мультиметодического подхода к конструктивной валидности, рассмотренного в главе 4. Самоотчет об эмоциях, хотя и дает ценную информацию, все же по разным причинам может оказаться неточным или неполным. Выражение лица, несомненно, может служить ключом к внутреннему переживанию, но одной только мимики недостаточно, чтобы судить об эмоциях. Дать однозначную интерпретацию физиологическим изменениям еще сложнее. Однако если все данные указывают в одном направлении, можно с гораздо большей уверенностью утверждать, что нам действительно удалось измерить интересующий нас конструкт. В обсуждаемой работе (Eisenberg et al., 1988) конвергенция данных была достигнута в двух аспектах. Во-первых, в том смысле, о котором мы только что говорили: все три показателя изменялись параллельно и в ожидаемом направлении, в соответствии с изменениями эмоциональной окраски стимулов. Вторая форма конвергенции —

Таблица 12.1

Тест эмпатии для детей и подростков Брайант

Утверждения		Ответ ^а
1	Мне грустно смотреть на девочку, которой не с кем поиграть	(+)
2	Целоваться и обниматься на людях — глупо	(-)
3	Мальчики, которые громко кричат, когда радуются, — глупы	(+)
4	Мне нравится наблюдать, как кто-то распаковывает подарок, даже если мне самому ничего не подарили	(+)
5	Мне хочется плакать, когда я вижу, что какой-то мальчик плачет	(+)
6	Я расстраиваюсь, когда вижу, что какая-то девочка сильно ударилась	(+)
7	Даже если я не знаю, почему кто-то смеется, мне хочется смеяться	(+)
8	Иногда, смотря телевизор, я плачу	(+)
9	Девочки, которые громко кричат, когда радуются, — глупы	(-)
10	Мне трудно понять, почему кто-то расстроился	(-)
11	Мне трудно смотреть, когда кто-то обижает животное	(+)
12	Мне грустно смотреть на мальчика, которому не с кем поиграть	(+)
13	Некоторые песни такие грустные, что когда я их слушаю, мне хочется плакать	(+)
14	Я расстраиваюсь, когда вижу, что какой-то мальчик сильно ударился	(+)
15	Взрослые иногда плачут, даже когда им вроде бы не о чем печалиться	(-)
16	Глупо относиться к собакам и кошкам так, будто они могут что-то чувствовать, как люди	(-)
17	Меня бесит, когда кто-нибудь из одноклассников постоянно притворяется, что ему нужна помощь учителя	(-)
18	Ребята, у которых нет друзей, наверное, просто не хотят, чтобы они у них были	(-)
19	Когда я вижу, как плачет какая-нибудь девочка, мне тоже хочется плакать	(+)
20	Мне кажется забавным, что некоторые люди плачут, когда смотрят грустный фильм или читают грустную книгу	(-)
21	Я могу спокойно есть свое печенье, даже если кто-то смотрит на меня, ожидая, что я ним поделюсь	(-)
22	Я не расстраиваюсь, когда вижу, как учитель наказывает моего одноклассника за непослушание	(-)

^а (+) указывает на то, что утвердительный ответ свидетельствует об эмпатии; (-) указывает на то, что отрицательный ответ свидетельствует об эмпатии.

Источник: *An Index of Empathy for Children and Adolescents*, В. К. Bryant, 1982, Child Development, 53, p. 416.

корреляционная: реакции детей по одному параметру коррелировали с их реакциями по двум другим параметрам. То есть грустное выражение лица обычно сопровождалось снижением частоты сердечных сокращений и сообщением о чувстве грусти — именно так, как и должно происходить в том случае, если все показатели отражают одно и то же эмоциональное состояние.

Когнитивный аспект

Последняя тема, которая будет рассмотрена в этом разделе, — когнитивный аспект нравственности. Теперь нас будет интересовать вопрос о том, как ребенок рассуждает о нравственности, и как изменяются его рассуждения в процессе развития когнитивных навыков.

Научно-исследовательские материалы, посвященные когнитивному аспекту нравственности, имеют два основных источника. Сначала обратимся к работам, начатым в 1920-х годах Пиаже и нашедшим свое продолжение в сотнях дальнейших исследований. Затем мы перейдем к рассмотрению современной исследовательской программы, инициатором которой стал Лоренс Колберг со своими коллегами.

Результаты исследований Пиаже, посвященных нравственности, объединены в одной книге, «Нравственные рассуждения ребенка» (*The Moral Judgment of the Child*, Piaget, 1932). Как отмечалось, исследования были проведены в 1920-х годах, что позволяет отнести их к ранним работам Пиаже. Из этого следует, что они не вполне согласуются с последующими работами, посвященными когнитивному развитию, которые мы рассматривали в главе 11. Тем не менее уже здесь намечены многие из основополагающих идей. Это и идея Пиаже о том, что нравственные рассуждения ребенка отражают не столько то, чему его научили родители или общество, сколько его собственный уровень когнитивного развития. Это и неизбежное следствие предыдущей мысли — идея о том, что изменения в нравственных суждениях являются результатом главным образом развития когнитивных навыков ребенка. И, с точки зрения методологии, как мы увидим, это подчеркивание необходимости использовать гибкий «клинический метод» тестирования как лучший способ прояснения представлений ребенка.

В своих исследованиях нравственных суждений Пиаже использовал два основных метода. Первый заключался в том, что детей расспрашивали о правилах разных игр, чаще всего об игре в стеклянные шарики. По мнению Пиаже, игры детей представляют собой своего рода микрокосм социального мира в целом, с собственными социальными нормами, закрепленными в процессе межличностного взаимодействия, санкциями за нарушение правил и т. д. Именно по этой причине, изучая игры, можно получить некоторое представление об уровне нравственного развития ребенка. Пиаже интересовала как способность ребенка придерживаться определенных правил, так и понимание им происхождения и сути правил¹. Беседы с ребенком, проводившиеся в гибкой манере клинического метода, должны были прояснить вопросы следующего типа: каковы правила игры; всегда ли правила игры были такими, какие они сейчас; можно ли изменить правила?

Вставка 12.2 иллюстрирует незначительную часть того, что удалось обнаружить Пиаже. Как можно заметить, представления детей о правилах в процессе развития определенно изменяются. Младшие дети склонны рассматривать правила как нечто неизменяемое, не поддающееся изменению — правила всегда были такими, какие они сейчас, они были сообщены Богом или отцом ребенка и не могут быть изменены. Старшие дети значительно яснее осознают то, что правила игр, если не

¹ Использование мужского рода в этом случае не имеет сексистской подоплеки. В 1920-х годах в Швейцарии в шарики играли только мальчики; поэтому испытуемыми в этой части исследований Пиаже тоже были только мальчики.

Вставка 12.2

Примеры ответов на вопросы Пиаже о правилах игры

ФОЛ (5) ... «Давным-давно, когда люди начинали строить город Невшатель, играли ли дети в шарики так же, как ты мне только что показал? — Да. — Всегда только так? — Да. — А как ты узнал о правилах? — Когда я был совсем маленький, мне показал их брат. А брату показал папа. — А как о них узнал твой папа? — Он просто знал. Никто ему не рассказывал. — Как он тогда узнал? — Никто ему не показывал!.. — Кто придумал игру в шарики? — Мой папа».

ШТОР (7) говорит, что дети играли в шарики еще до времен Ноева ковчега: «Как они играли? — Как и мы. — Как все началось? — Они купили несколько шариков. — Но как они научились? — Их папа научил». Штор изобретает новую игру для троих. Он полагает, что его друзьям понравится в нее играть, «но не всем. Не большим, не самым большим. — Почему? — Потому что это игра не для больших. — Это честная игра, как та, которую ты мне показал? — Нет. — Почему? — Потому что здесь не поровну. — А если бы кто-нибудь так играл, даже большие, это было бы честно? — Нет. — Почему? — Потому что здесь не поровну».

МОБ (12) ... «Все играют так, как ты мне только что показал? — Да. — Так играли и много лет назад? — Нет. — Почему? — Тогда использовали другие слова. — А правила? — Правила тоже были другие. Отец рассказывал, что они играли по-другому. — Но давным-давно люди играли по таким же правилам? — Не совсем по таким. — Что насчет правила не "бить" за один? — Я думаю, оно пришло позже. — А играли в шарики, когда твой дедушка был маленьким мальчиком? — Да. — Так же, как и сейчас? — Нет, по-другому. — ...Можно ли изменить правила? — Да. — А ты мог бы? — Да, я мог бы придумать новую игру. Однажды вечером мы играли дома и придумали новую [показывает ее нам]. — Эти правила такие же честные, как и в других играх? — Да. — А какая игра честнее, та которую ты показал мне в начале, или та, которую ты изобрел? — Они одинаково честные».

Источник: *The Moral Judgment of the Child* (p. 55, 60, 66-67), J. Piaget, 1932. New York: Free Press.

полностью, то частично условны и изменяемы. Этот сдвиг в представлениях об игре, с точки зрения Пиаже, отражает значительно более серьезный переход от «нравственности принуждения» или «нравственного реализма» к «нравственности сотрудничества» или «нравственному релятивизму». Вскоре мы рассмотрим еще несколько примеров этого перехода.

Второй метод изучения нравственных представлений, который использовал Пиаже, послужил отправной точкой для значительно большего количества исследований. Этот метод состоит в предъявлении нескольких рассказов, заключающих в себе некую нравственную дилемму, и выяснении мнения ребенка об этих рассказах. Самый известный пример — вопрос о том, как следует оценивать действия, принесшие вред: «объективно», то есть с точки зрения их материальных последствий, или «субъективно», то есть с точки зрения намерений, стоящих за этими действиями. Во вставке 12.3 представлены пять рассказов, которые использовал Пиаже для изучения этого вопроса; во вставке 12.4 приведены примеры ответов детей. Как можно заметить, в процессе развития происходит переход от «объективного» ориентирования на последствия к более зрелому «субъективному» учету намерений. Получить представление о некоторых других вопросах, которые исследовал Пиаже, можно, ознакомившись с двумя примерами рассказов и соответствующих ответов, приведённых во вставке 12.5. Первый пример касается вопроса о том, почему нельзя лгать, а второй пример связан с понятием «имманентной справедливости» или веры в неизбежность наказания за плохое поведение.

Вставка 12.3

Рассказы, использовавшиеся Пиаже для изучения объективной/ субъективной ответственности

I. А. Маленький мальчик по имени Джон сидел в своей комнате. Его позвали обедать. Он пошел в столовую, но за дверью стоял стул, а на стуле — поднос с 15 чашками. Джон не мог знать, что он там стоит. Он вошел, дверь сбила поднос, и все чашки разбились!

В. Жил на свете маленький мальчик по имени Генри. Однажды, когда его матери не было дома, он попытался взять из буфета банку с джемом. Он взобрался на стул, но не смог достать до банки. Пытаясь дотянуться до нее, он уронил чашку. Чашка упала и разбилась.

II. А. Жил на свете мальчик по имени Джулиан. Его отца не было дома, и Джулиану подумалось, что будет очень интересно поиграть с чернильницей отца. Он начал играть с пером, но поставил небольшую кляксу на скатерть.

В. Однажды маленький мальчик по имени Агустус заметил, что чернильница его отца пуста. Как-то раз, когда отца не было дома, он решил наполнить чернильницу, чтобы помочь отцу. Но, открывая флакончик с чернилами, Агустус поставил на скатерть большую кляксу.

III. А. Жила на свете маленькая девочка по имени Мэри. Она захотела сделать матери подарок и что-нибудь для нее сшить. Но, не умея как следует пользоваться ножницами, она сделала на своем платье огромную дырку.

В. Однажды, когда матери не было дома, маленькая девочка по имени Маргарет пошла и взяла мамини ножницы. Не умея правильно с ними обращаться, она сделала небольшую дырочку на своем платье.

IV. А. Как-то раз Альфред встретил своего маленького и очень бедного друга. Друг рассказал ему, что он еще не пообедал, потому что в доме нет ничего поесть. Альфред пошел в булочную и, поскольку у него не было денег, дождался, когда продавец отвернулся, и украл булочку. Он выбежал из магазина и отдал булочку своему другу.

В. Генриетта зашла в магазин и увидела на столе красивую ленту. Она подумала, что эта лента очень хорошо подошла бы к ее платью. Когда продавщица отвернулась, Генриетта украла ленточку и тут же убежала.

V. А. У Альбертины была подруга, которая держала в клетке птицу. Альбертине казалось, что птице очень грустно, и она не раз просила подругу, чтобы та ее отпустила. Но подруга не соглашалась. Однажды, когда подруги не было, Альбертина пришла и украла птицу. Она выпустила ее на волю, а клетку спрята-ла на чердаке, чтобы в нее больше не попала ни одна птица.

В. Однажды, когда матери не было дома, Джулия украла из буфета несколько конфет, спрятала их, а потом тайком съела.

Источник: *The Moral Judgment of the Child* (p. 122–123), J. Piaget, 1932. New York: Free Press.

Проанализируем рассказы из вставки 12.3 и вопрос «объективной/субъективной» нравственности. Многое из методики Пиаже заслужило критику исследователей, и прежде чем читать дальше, полезно поразмышлять об этих недостатках. Фактически, у меня есть возможность перечислить лишь некоторые из модификаций методики Пиаже. Одни специалисты заявляют о том, что порядок событий в рассказах Пиаже, где информация о последствиях всегда стоит на последнем месте, может заставить детей уделять основное внимание последствиям, а не намерениям героя; поэтому производилось изменение порядка событий в рассказах с использованием либо внутрисубъектного, либо межсубъектного плана исследования (например, Moran & McCullers, 1984). Другие специалисты подчеркивают трудно-

Вставка 12.4

Примеры ответов на рассказы из вставки 12.3

ГЕО (6): «Ты понял эти рассказы? — Да. — Что сделал первый мальчик? — Он разбил двенадцать чашек. — А второй? — Он задел и разбил одну чашку. — Почему первый мальчик разбил чашки? — Потому что их сбила дверь. — А второй? — Он был неуклюжим. Когда он тянулся за джемом, чашка упала. — Кто-нибудь из мальчиков виноват больше? — Первый, потому что он разбил двенадцать чашек. — Если бы ты был папой, кого из них ты бы наказал строже? — Того, который разбил двенадцать чашек».

КОНСТ (7) ...правильно пересказывает историю с чернильным пятном: «Маленький мальчик видит, что чернильница его отца пустая. Он берет флакончик с чернилами, но ставит большую кляксу. — А другой? — Жил такой мальчик, который все трогал руками. Он взял чернила и поставил маленькую кляксу. — Мальчики одинаково провинились или нет? — Нет. — А кто из них виноват больше? — Тот, который поставил большую кляксу. — Почему? — Потому что она большая. — А почему он поставил большую кляксу? — Он хотел помочь. — А почему второй поставил маленькую кляксу? — Потому что он все трогал руками. Он поставил маленькую кляксу. — И кто же из них виноват больше? — Тот, который поставил большую кляксу».

ГРОС (9): «Что сделал первый? — Он разбил пятнадцать чашек, когда открывал дверь. — А второй? — Он разбил одну чашку, когда пытался достать джем. — Как ты думаешь, в какой из этих нелепостей — большее непослушание? — Та, где мальчик пытался удержать чашку [нелепее], потому что другой не видел [что за дверью стояли чашки]. Он видел, что делал. — Сколько он разбил? — Одну чашку. — А второй? — Пятнадцать. — Кого тогда нужно строже наказать? — Того, кто разбил одну чашку. — Почему? — Он нарочно это сделал. Если бы он не полез за джемом, этого бы не случилось».

НАСС (10): Больше провинился тот, кто хотел достать банку с джемом. — Имеет ли значение то, что другой разбил больше чашек? — Нет, потому что тот, который разбил пятнадцать чашек, сделал это не нарочно».

Источник: *The Moral Judgment of the Child* (p. 124–125, 126, 129–130), J. Piaget, 1932. New York: Free Press.

сти, с которыми могут столкнуться маленькие дети, пытаюсь понять и запомнить информацию, содержащуюся в воспринимаемых на слух парах историй. Задача упрощалась предъявлением рассказов по одному с последующим сравнением суждений ребенка о непослушании героев разных рассказов (например, Berg-Cross, 1975). Другие исследователи заменяли устную форму представления видеосюжетами о поведении, относительно которого нужно было вынести некое суждение (например, Chandler, Greenspan, & Barenboim, 1973). Ряд исследователей (например, Nelson-LeGall, 1985) отмечают, что в рассказах Пиаже не всегда четко разделяется мотив (во благо или во вред?) и умысел (преднамеренно или случайно?); по этой причине производились модификации, в которых эти два аспекта разво-дились. Наконец, наверное, самое важное изменение, которое было внесено в последующих исследованиях, связано с попыткой преодолеть главное смешение факторов в рассказах Пиаже. Поскольку мотив и степень ущерба в большинстве рассказов Пиаже изменяются параллельно, невозможно определить, какую именно информацию способен использовать ребенок.

На основе указанных выше исследований можно сделать три главных вывода (см. также Langford, 1995). Эти выводы перекликаются с выводами, сделанными на основе исследований, направленных на проверку теории Пиаже о когнитивном развитии. Во-первых, использование методов Пиаже приводит к некоторой недооценке способностей маленького ребенка; при использовании видеоизмененных

Вставка 12.5

Примеры рассказов Пиаже и ответов на вопросы о лжи и имманентной справедливости

А. «Маленький мальчик [или маленькая девочка] гулял по улице и встретил большую собаку, которая его сильно напугала. Он пришел домой и сказал матери, что он видел собаку величиной с корову».

В. «Один ребенок пришел домой со школы и сказал своей матери, что получил хорошие отметки, но это была неправда; учитель вообще не ставил ему никаких оценок, ни плохих, ни хороших. Его мать была очень довольна и сделала ему подарок».

ФЕЛ (6): правильно пересказывает обе истории: «Кто из этих двух детей провинился больше? — Девочка, которая сказала, что видела собаку величиной с корову. — А почему она провинилась больше? — Потому что такого не бывает. — А ее мать поверила ей? — Нет, потому что нет таких [собак, которые были бы величиной с корову]. — Почему она так сказала? — Чтобы преувеличить. — А почему другая солгала? — Потому что она хотела, чтобы все поверили, что она получила хорошую отметку. — Мать поверила ей? — Да. — Если бы ты была матерью, кого бы ты наказала строже? — Ту, которая рассказала про собаку, это совершенное вранье, поэтому она виновата больше».

ЭРЛ (10): «Виноват больше тот, кто обманул мать, сказав, что учитель был им доволен. — Почему он виноват больше? — Потому что мать отлично знает, что собак величиной с корову не бывает. Но она верит мальчику, который сказал, что учитель был им доволен. — Почему ребенок сказал, что собака была величиной с корову? — Чтобы ему поверили. Чтобы подшутить. — А почему другой сказал, что учитель был им доволен? — Потому что он плохо сделал задание, которое дал учитель. — Это шутка? — Нет, это ложь. — А шутка это то же самое, что ложь? — Ложь хуже, потому что она вреднее».

Как-то раз двое детей воровали яблоки в саду. Но тут вдруг пришел полицейский, и дети кинулись бежать. Одного из них поймали. Другой побежал домой окольным путем, переходил через реку по разбитому мосту и упал в воду. Как ты думаешь, если бы он не украл яблоки и не переходил реку по разбитому мосту, упал бы он все равно в воду?

ПЭЛ (7): ...«Что ты об этом думаешь? — Это справедливо. Так ему и надо. — Почему? — Потому что нельзя воровать. — Если бы он не воровал, упал бы он тогда в воду? — Нет. — Почему? — Потому что он не сделал бы ничего плохого. — Почему он упал? — Это ему в наказание».

ФРЭН (13): ...«Если бы он не крал яблоки, упал бы он тогда в воду? — Да. Если бы мост провалился, он провалился бы в любом случае, потому что он был разбит».

Источник: *The Moral Judgment of the Child* (p. 148, 150–151, 157–158, 252, 253–254, 255), J. Piaget, 1932. New York: Free Press.

процедур ответы детей нередко свидетельствуют о более высоком нравственном уровне. Во-вторых, нравственные суждения представляют собой более сложный феномен, чем полагал Пиаже, определяющийся значительно большим количеством факторов; на ответ ребенка могут влиять самые разнообразные переменные. Наконец, несмотря на все недочеты работ Пиаже, нет никаких сомнений в том, что они имеют огромное значение, и не только в связи с конкретными результатами, но и в связи с тем, что в них впервые в качестве области исследования рассматривался когнитивный аспект нравственности. Обратимся теперь к главной современной концепции нравственного развития, к концепции Лоренса Колберга.

Как и Пиаже, Колберг в основу своего подхода положил реакцию испытуемого на гипотетические нравственные дилеммы. Однако Колберга интересуют нравственные рассуждения более высокого уровня, чем те, которые изучал Пиаже; и

Вставка 12.6

Примеры нравственных дилемм Колберга

Дилемма III: Одна женщина, жившая в европейской стране, умирала от тяжелой формы рака. Было только одно лекарство, которое, по мнению врачей, могло бы ее спасти — особая форма радия, которую недавно в том же городе открыл один фармаколог. Изготовление этого лекарства обходилось очень дорого, но фармаколог запрашивал за него в десять раз больше, чем стоило его изготовление. Он платил за радий 200\$, а за небольшую дозу лекарства просил 2000\$. Муж больной женщины, Хайнц, обращался ко всем, кого он знал, чтобы занять денег, но ему удалось собрать только 1000\$ — половину необходимой суммы. Он сказал фармакологу, что его жена умирает, и попросил его отпустить лекарство по более низкой цене или в кредит. Но фармаколог ответил: «Нет, я открыл новое лекарство и хочу сделать на этом деньги». Хайнц был в отчаянии и стал обдумывать, не пробраться ли ему на склад и не украсть ли это лекарство. Следует ли Хайнцу красть лекарство? Почему?

Дилемма V: На войне в Корее один отряд морских пехотинцев нес большие потери и был вынужден отступать. Отряд пересек мост через реку, однако на другой стороне реки оставались большие силы врага. Если бы кто-нибудь вернулся и взорвал мост, имея позиционное преимущество, оставшиеся сумели бы уйти. Но тот, кто взялся бы взорвать мост, наверняка бы погиб; шансы остаться в живых составляли 1:4. Капитан знал лучше всех в отряде, как правильно вести отступление. Он спросил, не вызовется ли кто-нибудь добровольцем, но добровольцев не оказалось. Если бы он пошел сам, многие, вероятно, не смогли бы вернуться, потому что он был единственным, кто умел вести отступление. Следует ли капитану поручить выполнение этого очень опасного задания кому-то из отряда или пойти самому? Почему?

Дилемма VIII: В одной европейской стране бедняк по имени Вальжан не мог найти себе работу; не могли найти работу и его брат и сестра. Не имея денег, Вальжан воровал еду и лекарства, в которых они нуждались. Его поймали и посадили в тюрьму на шесть лет. По прошествии нескольких лет, он сбежал из тюрьмы и под новым именем отправился жить в другую часть страны. Он накопил денег и построил большую фабрику. Он платил рабочим очень высокую зарплату и использовал значительную часть прибыли на постройку больницы для тех, кто не может позволить себе достойное медицинское обслуживание. Прошло 20 лет, и вот однажды некий портной узнал во владельце фабрики беглого преступника Вальжана, которого полиция искала в его родном городе. Следует ли портному заявлять о Вальжане в полицию? Правильно было бы никому не рассказывать о происшедшем или нет? Почему?

Источник: A Longitudinal Study of Moral Judgment, A. Colby, L. Kohlberg, J. Gibbs, & M. Lieberman, 1983, Monographs of the Society for Research in Child Development, 48, p. 77, 82, 83.

дилеммы, которые он предлагает, соответственно значительно сложнее. Всего дилемм Колберга девять. Во вставке 12.6 представлены три из них, включая наиболее известную и часто приводимую в качестве примера — историю Хайнца.

Первое, что требуется от испытуемого, ознакомившегося с дилеммой, — вынести положительное или отрицательное суждение о нравственности поведения героя рассказа (например: «Должен ли был Хайнц красть лекарство?», «Должен был капитан приказать рядовому выполнить задание, или должен был пойти сам?»). Однако по-настоящему интересно обоснование этого суждения. Сначала на подобное рассуждение испытуемого должен вызвать вопрос «почему»; затем экспериментатор задает всевозможные полустандартные вопросы, стараясь уточнить мнение испытуемого. Для оценки ответов и стадий нравственного развития, которым они соответствуют, решающим является именно ход рассуждения, а не само по себе вынесенное положительное или отрицательное суждение о нравственности.

Стадии нравственного рассуждения по Колбергу

Уровни и стадии	Описание	Доводы за	Доводы против
<i>Доконвенционный уровень</i>			
Стадия 1. Ориентированность на наказание/послушание	Соблюдение норм для избежания наказания. Послушание ради самого послушания и непричинение физического вреда лицам и имуществу	Ему нужно украсть это лекарство. Это не такой уж плохой поступок. Лекарство, которое бы он украл, стоит \$ 200; он ведь украл бы лекарство не на \$ 2000	Хайнцу не нужно красть это лекарство; он должен его купить. Если он его украдет, то может попасть в тюрьму, и тогда ему все равно придется вернуть лекарство
Стадия 2. Ориентированность на выгодный обмен	Соблюдение норм только тогда, когда это приносит выгоду; действия только в собственных интересах, удовлетворение собственных потребностей и признание права других делать то же самое. Правильно — то, что по честному: равноценный обмен, сделка, соглашение	Хайнцу нужно украсть лекарство, чтобы спасти свою жену. Его могут отправить в тюрьму, но у него останется жена	Ему не следует красть. Фармаколог не делает что-то предосудительное; он просто хочет получить прибыль. Для этого люди и занимаются бизнесом — чтобы заработать денег
<i>Конвенционный уровень</i>			
Стадия 3. Ориентированность на «хорошее» поведение	Поведение в соответствии с ожиданиями близких людей или в соответствии с тем, что люди вообще ожидают от человека в роли сына, брата, друга и т. д. Быть «хорошим» важно и означает действовать из благородных побуждений, проявлять заботу о других, поддерживать отношения на основе доверия, преданности, уважения и благодарности	Если бы я был Хайнцом, то украл бы лекарство для своей жены. Нельзя навесить ценник на любовь, ее нельзя купить. Жизнь тоже не имеет цены	Ему не следует красть. Если его жена умрет, он не будет виноват. Это не потому, что он бессердечен или не слишком сильно любит свою жену, чтобы сделать для нее все возможное. Это фармаколог бессердечный эгоист

Уровни и стадии	Описание	Доводы за	Доводы против
Стадия 4. Ориентированность на поддержание порядка	Выполнение своих обязанностей и обязательств. Нарушение законов возможно только в крайних случаях, когда они вступают в противоречие с социальными обязательствами. Долг человека — приносить пользу обществу, группе или социальному институту	Вступая в брак, ты клянешься любить свою жену и заботиться о ней. Брак — это не только любовь, это еще и обязательство. Как юридический документ	Нет ничего удивительного в том, что Хайнц хочет спасти свою жену; но воровать нельзя никогда. Он знает, что, воруя, отнимает ценное лекарство у человека, который его изготовил
Постконвенционный уровень			
Стадия 5. Ориентирование на общественный договор	Признание того, что люди исповедуют разные ценности и придерживаются разных мнений, что большинство ценностей и норм относительны. Эти относительные нормы следует соблюдать в интересах справедливости и потому, что это — общественный договор. Однако некоторые ценности, такие как жизнь и свобода, должны охраняться в любом обществе, вне зависимости от мнения большинства	Для этих обстоятельств нет закона. Украсть лекарство было бы неправильно, но оправдано	Здесь нельзя свалить всю вину на кого-то одного, но чрезвычайные обстоятельства не дают права брать закон в свои руки. Нельзя, чтобы каждый отчаявшийся начал воровать. Цель может быть и благородна, но цели не оправдывают средства
Стадия 6. Ориентирование на универсальные этические принципы	Следование самостоятельно выбранным этическим принципам. Конкретные законы и социальные соглашения имеют силу потому, что они основаны на подобных принципах. Если законы нарушают эти принципы, следует действовать в соответствии с принципами, поскольку они универсальны: равноправие и уважение достоинства человека как самостоятельной личности	Эта ситуация заставляет его выбирать между возможностью украсть и возможностью дать своей жене умереть. В ситуации, где необходимо сделать такой выбор, с точки зрения нравственности, правильно решиться на воровство. Он должен действовать, руководствуясь принципами сохранения жизни и уважения к праву на жизнь	Перед Хайнцем стоит вопрос, думать ли о жизнях других людей, которые нуждаются в этом лекарстве не меньше его жены. Хайнц должен действовать не в соответствии с чувствами к своей жене, а уважая жизнь каждого человека, кому необходимо это лекарство

Источник: *The Stage Sequence: The Cognitive-Development Approach to Socialization* (p. 379–380), L. Kolberg. In D. A. Goslin (Ed.), *Handbook of Socialization Theory and Research* (p. 347–480), 1969, Chicago: Rand McNally.

Чаще всего при изложении теории Колберга выделяют упорядоченную последовательность из шести стадий развития¹. Эти стадии, в свою очередь, можно объединить в три уровня развития. В табл. 12.2 перечислены стадии и соответствующие уровни, а также примеры ответов по истории Хайнца. Как явствует из таблицы, в процессе развития происходит постепенное нарастание сложности рассуждений: ребенок переходит от ориентирования на награду и наказание к учету ожиданий и норм общества, а затем — к формулированию внутренних принципов и представлений о совести. Заметьте, что высший уровень достигается, как правило, не ранее подросткового возраста и наблюдается далеко не у всех взрослых.

Мы еще не затрагивали вопрос о том, как оцениваются ответы испытуемых, и как их относят к определенной стадии нравственного развития. По ряду причин объяснить это в нескольких словах затруднительно. Во-первых, система оценки Колберга весьма сложна; это, наверное, самая сложная система оценки в психологической литературе. Во-вторых, с момента начала Колбергом своих работ по проблеме нравственного развития 40 лет назад система оценки претерпела существенные изменения. Сегодняшняя система оценки, «Standard Issue» (Colby et al., 1987), — третья пересмотренная версия системы Колберга и его коллег. Существенность изменений демонстрирует тот факт, что коэффициент корреляции между результатами, полученными с помощью современной версии, и результатами, полученными при использовании оригинальной версии, составляет всего 0,39.

Конечно, сложность системы оценки и необходимость внесения изменений имеют свои причины. В целом, сотрудники исследовательской группы Колберга при неоднократном пересмотре пытались привести систему оценки в максимальное соответствие с положениями теории Колберга о том, как развивается нравственное суждение. В частности, со временем оценка становилась все менее связанной со специфическим содержанием ответа ребенка и все более ориентированной, как и в теории Колберга, на уровень и структуру рассуждения. Внесение изменений имеет и психометрические причины. Первые системы оценок Колберга были субъективны и сложны в практическом применении, поэтому большие сомнения вызывала и их надежность (Kurtines & Greif, 1974). Последняя версия, хотя и достаточно трудна для использования, все же обладает неплохой надежностью, основанной на общности оценок нескольких специалистов, и приемлемой ретестовой надежностью (Colby & Kolberg, 1987).

Хотя сложность системы оценки Колберга и оправдана, она тем не менее затрудняет проведение исследования. Исследователь, которого заинтересовала данная тема, не может просто пойти и провести «колберговское исследование» так же, как он мог бы провести «пиagetовское исследование». Для проведения исследования по системе Колберга необходимо получить доступ к специальным материалам и пройти подготовку в специальной группе. В действительности, даже чтобы оценить подобные исследования, может потребоваться доступ к первоисточникам и к информации о системе, использовавшейся при обработке данных. Как след-

¹ В последних работах Колберга (например, Colby, Kohlberg, Gibbs, & Lieberman, 1983) стадии 5 и 6 были объединены, то есть 6-ступенчатая модель превратилась в 5-ступенчатую.

ствие, «колберговские исследования» проводят в основном сам Колберг, его ученики и сотрудники его исследовательской группы. Это положение дел затрудняет осуществление независимой проверки теории.

Учитывая все эти проблемы, следует отметить, что было разработано две альтернативы стандартному подходу Колберга. Первая альтернатива — «Показатель социоморальной рефлексии» (*Sociomoral Reflection Measure* или *SRM*) (Gibbs, Widman, & Colby, 1982). *SRM* включает в себя те же нравственные дилеммы и предназначен для выявления тех же типов нравственных рассуждений, что и методика Колберга. Однако в *SRM* дилеммы представлены в письменном виде, и отвечает на них испытуемый также в письменном виде. То есть *SRM* можно использовать для группового тестирования, а предъявление и обработка результатов значительно упрощаются. Не так давно была разработана сокращенная версия «Показателя социоморальной рефлексии» — *SRM-SF* (Gibbs, Basinger, & Fuller, 1992). Это результат дальнейшего упрощения с использованием прямой формулировки нравственных проблем, а не полного изложения дилемм (например, «Представьте, что вы дали обещание другу. Насколько важно выполнять обещания?...»). Гиббс и коллеги (Gibbs, Basinger, & Fuller, 1992) сообщают о впечатляющем уровне соответствия оценок стадии нравственных рассуждений, полученных с помощью их методики и стандартной методики Колберга. Они также приводят обширные данные, доказывающие надежность и валидность *SRM* и *SRM-SF*.

Вторая из важнейших альтернатив методики Колберга — «Тест определяющих проблем» (*Defining Issues Test* или *DIT*), разработанный Джеймсом Рестом (James Rest, 1979). В *DIT* используются те же нравственные дилеммы, что и в методике Колберга. Однако в *DIT* вслед за дилеммой идет список из 12 проблем, который испытуемый должен оценить по уровню значимости при принятии решения в описанной ситуации. Например, после истории Хайнца стоят вопросы следующего типа: «Следует соблюдать законы общества или нет?», «Какими ценностями руководствуются люди, взаимодействуя друг с другом?» О стадии нравственного развития судят по оценке испытуемым важности этих вопросов. Таким образом, *DIT* дает показатель понимания и оценки, основанный на выборе ответов, и тем самым отличается от методики Колберга, в которой оцениваются более или менее свободные рассуждения испытуемых. Коэффициент корреляции между результатами по этим двум тестам умеренный, но не абсолютный; обычно он находится в пределах 0,6–0,7. С точки зрения методологии, *DIT* имеет несколько преимуществ: на его выполнение уходит меньше времени, чем на тест Колберга; его результаты в меньшей степени зависят от способности вербально выражать свои мысли и подлежат объективной оценке.

Перейдем теперь от самого метода к краткому рассмотрению вопросов, поднятых работами Колберга. Множество вопросов вызвали как исследовательский интерес, так и разногласия, включая проблему связи между нравственными суждениями и нравственным поведением (например, Blasi, 1980), гипотезу о различии между полами в нравственных суждениях (например, Gilligan, 1982) и возможности создания программ нравственного воспитания, основанных на теории Колберга (например, Sockett, 1992). Здесь мы сконцентрируемся на вопросе, имеющем

непосредственное отношение к теории: на гипотезе о том, что в процессе развития нравственные суждения проходят ряд стадий. Как и теория стадийности Пиаже, рассмотренная в главе 11, эта теория позволяет сделать два предположения о взаимосвязи развития отдельных навыков. Первое, что можно ожидать, это согласованность или синхрония: если имеет смысл говорить, что дети находятся на некоторой стадии, их суждения должны соответствовать этой стадии. Второе, что можно ожидать, это инвариантная последовательность: низшие стадии должны всегда предшествовать высшим, и ни один ребенок не может ни перескочить через стадию, ни вернуться назад.

Результаты исследований в целом подтверждают эти предположения. Для проверки гипотезы о синхронии необходимо использовать внутрисубъектный подход, при котором одни и те же испытуемые дают ответы на все дилеммы, что позволяет сравнить их рассуждения при выполнении разных заданий. В подобных исследованиях обнаруживается существенная, но не полная согласованность ответов по уровню нравственности. Уокер, де Врие и Треветан (Walker, deVries, & Trevethan, 1987), к примеру, изучали ответы как на стандартные дилеммы Колберга, так и на «реальные» дилеммы, которые предлагали сами испытуемые. Они обнаружили, что 62 % выборки находилось по обоим параметрам на одной стадии и что в 90 % случаев стадии либо полностью совпадали, либо были смежными. Аналогичные результаты были получены при сравнении ответов на несколько стандартных дилемм. Обычно преобладающая часть ответов респондентов, 65 %–70 %, относится к какой-либо одной модальной стадии, а большинство остальных ответов можно отнести к стадии либо следующей за модальной, либо предшествующей ей. Расхождения в две и более стадий крайне редки (Walker, 1988).

Для проверки гипотезы об инвариантной последовательности необходимо проведение лонгитюдного исследования. Первое подобное исследование инициировал сам Колберг: его диссертационная работа переросла в 20-летний проект; при этом первый раз испытуемых тестировали в конце 1950-х годов, когда им было 10–16 лет, а затем, в течение следующих 20 лет, они проходили тестирование пять раз с интервалом в 3–4 года. В итоговой публикации, содержащей выводы, полученные в этом исследовании (Colby, Kolberg, Gibbs, & Lieberman, 1983), указывается на отсутствие случаев пропуска каких-либо стадий и на наличие лишь незначительного количества случаев снижения показателей уровня нравственности, что объясняется скорее погрешностью измерения, а не истинным регрессом. Однако следует отметить, что признаки регресса в данных первого лонгитюдного исследования послужили одним из оснований для пересмотра системы оценки; то есть регресс был исключен помещением кажущихся незрелыми ответов на более высокие уровни. Хотя этот подход оправдан (Colby et al., 1983), у скептиков он вызывает сомнение в том, насколько вообще возможна проверка положения об инвариантной последовательности. С другой стороны, последние лонгитюдные исследования, основанные на современном варианте системы оценки, это положение неизменно подтверждают (Walker, 1989; Walker & Taylor, 1991).

Проверка положения о стадийности осуществляется и с помощью кросс-культурных исследований. Следуя теории Колберга, с ее акцентом на базовом когнитивно-структуральном компоненте нравственности, можно предположить наличие

существенного сходства в нравственном развитии в разнообразных культурных условиях. Теория допускает различие в темпах развития, в конечном уровне и в специфическом содержании некоторых ответов. Однако в ней подчеркивается, что во всех культурах можно выявить одни и те же базовые стадии, которые преодолеваются представителями всех культур в одном и том же порядке. В целом результаты кросс-культурных исследований, которые сегодня охватывают впечатляющее количество культур и впечатляющее их разнообразие, подтверждают это теоретическое предположение (Edwards, 1986; Snarey, 1985). С некоторыми методологическими коррективами, стадии, о которых говорит Колберг, обнаруживаются во всех изучаемых культурах, возрастные тренды согласуются с представлением о переходе с низших стадий на высшие, а данные, полученные в ряде лонгитюдных исследований в других культурах, подтверждают положение об инвариантной последовательности. Однако из полученных данных также следует, что стадии выше четвертой среди представителей незападных культур очень редки. На сегодняшний момент неясно, отражает этот факт истинные межкультурные различия в уровне нравственных рассуждений, или то, что Колбергу не удалось учесть формы нравственного мышления, которые имеют значение в культурах, отличных от нашей.

Последнее, о чем следует упомянуть в этом разделе, имеет некоторое отношение к сказанному выше. Даже в рамках нашей культуры теории Пиаже и Колберга, какими бы информативными они ни были, не исчерпывают область когнитивного аспекта нравственности. В последние годы ведется разработка ряда исследовательских программ, в целом соответствующих когнитивно-структуралистскому духу теории Пиаже и Колберга, но воплощающих в себе попытку исследовать формы рассуждения, которым эти ученые не уделили должного внимания. В этом отношении стоят упоминания работы, посвященные рассуждениям детей о просоциальном поведении (например, Eisenberg, 1982), рассуждениям о справедливости распределения (например, Damon, 1980), и пониманию детьми социальных условностей и разницы между нравственностью и условностями (например, Smetana, 1993; Turiel, 1983). Туриель (Turiel, in press) делает обзор этих работ, а также последних исследований, проведенных в традиции Пиаже и Колберга.

Я-концепция

Как и нравственность, Я-концепция — большая, многогранная тема. Так же как и изложение темы нравственности, изложение темы Я-концепция будет избирательным, с соблюдением общепринятого разделения на подтемы. В данном случае производится деление на когнитивный компонент Я-концепции и на оценочный компонент — на то, что дети думают или знают о себе, и на то, какого они о себе мнения. Первое называют самосознанием, второе — самооценкой.

Самосознание

Исследования развития самосознания у детей начинаются с постановки базового вопроса: Когда ребенок впервые осознает, что он обладает собственным Я? И тут же неизбежно встает методологический вопрос: какие данные мы могли бы использовать для выяснения того, что знает о себе не владеющий речью младенец или

плохо владеющий речью 2–3-летний ребенок? Это, как вам, наверное, понятно, лишь новый вариант вопроса, с которым мы уже сталкивались, рассматривая разнообразные аспекты развития младенца.

При рождении и в первые недели жизни малыши не делают ничего, что наводило бы на мысль — по крайней мере, психологов — о том, что у них имеются какие-то представления о самих себе. В действительности, одним из положений многих теорий, в частности теории Пиаже, является положение о том, что дети вступают в жизнь, будучи совершенно неспособными выделить себя из внешнего мира. Однако постепенно начинает вырисовываться поведение, свидетельствующее об элементарных зачатках осознания самого себя. Чаще всего в этой связи упоминают разнообразные навыки, относящиеся к категории **персонального влияния**: понимания ребенком того, что он может являться причиной неких событий во внешнем мире. Несколько недель спустя после рождения младенцы производят действия, подразумевающие наличие у них ощущения собственного влияния, а к середине первого года жизни они демонстрируют понимание своей активности уже весьма недвусмысленно. Об этом, к примеру, свидетельствуют результаты исследований Пиаже; фактически, сенсомоторное развитие по большей части заключается в освоении младенцем все более эффективных способов влияния на среду, направленных на получение желаемого. Об этом также свидетельствуют иллюстрируемые рис. 11.2 исследования Рови-Колльер (Rovee-Collier & Bhatt, 1993), в которых младенцы узнают, что они могут контролировать движение подвешенных над их кроватками игрушек. Об этом свидетельствует и ряд других феноменов, выявленных в исследованиях социального развития в раннем возрасте, главный вывод из которых состоит в том, что с самого раннего возраста малыши становятся активными участниками социального взаимодействия, демонстрируя поведение (плач, жесты и т. д.), функция которых заключается во влиянии на поведение другого человека.

Персональное посредничество обычно рассматривается как субъективный аспект Я — то есть как осознание ребенком своей активной роли в генерировании некоторого опыта. Интерес представляет также и объективный аспект Я как объекта созерцания — то, что ребенок думает или знает о себе. Здесь также первые признаки неизбежно носят невербальный характер — это разные формы поведения, указывающие на то, что ребенок начинает понимать, кто он такой. Главным признаком считается **визуальное самоузнавание**. На каком этапе развития младенец впервые узнает свое изображение? Допустим, мы ставим перед малышом зеркало: поймет ли он, что странное существо, которое он видит, — это на самом деле он сам? В случае годовалого ребенка ответ на этот вопрос, скорее всего отрицательный. Дети в возрасте до года нередко позитивно реагируют на зеркальный образ: улыбаются, тянутся ручками к лицу, которое видят; однако это поведение ничуть не отличается от реакции ребенка на других малышей. Лишь где-то на втором году жизни ребенка мы начинаем замечать первые попытки установить связь между собой и образом. Наиболее отчетливо эти признаки проявляются при проведении методики, известной как тест с помадой.

Тест с помадой независимо друг от друга разработали два исследователя: Амстердам (Amsterdam, 1972) — для изучения младенцев; и Гэллуп (Gallup, 1970) —

для изучения обезьян. Тест начинается с того, что мать незаметно для ребенка ставит ему помадой точку на нос или на лоб. Затем перед ребенком ставят зеркало и регистрируют реакцию малыша. Вопрос, естественно, заключается в том, будет ли ребенок прикасаться к следу помады. Если реакция будет именно таковой, это будет достаточно убедительным свидетельством того, что ребенок осознает связь между отражением в зеркале и самим собой. До 15-месячного возраста подобный интерес к себе — редкое явление. Прикосновение к лицу и другие реакции, имеющие отношение к самоузнаванию (например, замешательство ребенка, заметившего след от помады), начинают появляться в период между 15 и 18 месяцами жизни и становятся совершенно отчетливыми к 24 месяцу жизни (Povinelli, 1995). Таким образом, к этому возрасту, несомненно, имеет место самоузнавание¹.

Другие признаки самоузнавания наблюдаются либо в то же время, либо вскоре после появления адекватной реакции в тесте с помадой (Stipek, Gralinski, & Kopp, 1990). Двухлетние дети обычно в состоянии узнать себя не только в зеркале, но и на фотографии или на видеозаписи. Большинство 2-летних детей уже знают свое имя и могут правильно его использовать, упоминая самих себя. Кроме того, начинают использоваться местоимения первого лица — я, мне, меня, мое.

С развитием у ребенка речи все чаще начинают прибегать к вербальным методикам оценки Я-концепции. Основной подход при работе со старшими детьми — самоописание: ребенка просят описать, каков он (Damon & Hart, 1988; Harter, 1988). Обычно проведение таких методик начинается с общего вопроса, не ограничивающего высказывания детей, — к примеру: «Расскажи, какой ты» или «Какой ты человек?» Затем, как правило, следуют более конкретные вопросы, форма которых зависит от того, что именно интересует исследователя; исследуется такая проблематика, как Я-концепция в разных условиях и разных отношениях (например: «Расскажи, какой ты со своей семьей?» или «Расскажи, какой ты со своими друзьями») и Я-концепция в процессе развития (например: «Как ты думаешь, через пять лет ты будешь таким же или другим?», «Ты вообще изменяешься с каждым годом?»).

Как следует из вышесказанного, при помощи самоописания можно изучать ряд вопросов. Главный из них, однако, связан с развитием Я-концепции: как изменяются рассуждения ребенка о самом себе в процессе развития? С возрастом действительно происходят изменения, и в целом они соответствуют тому, что можно было бы ожидать, имея представления об изменениях, происходящих в процессе когнитивного развития. В табл. 12.3 приводятся примеры обычной последовательности этапов развития. Как можно заметить, в своих самоописаниях дошкольники сосредоточиваются в основном на внешних, ситуативных характеристиках: физических атрибутах, вещах, которыми они владеют, предпочтениях. К периоду среднего детства акцент смещается на менее осязаемые характеристики, такие как эмоции, поведенческие особенности и склонности. Наконец, в подростковом возрасте появляются абстрактные характеристики, характеристики еще более высокого по-

¹ Как отмечалось, самоузнавание изучалось также на обезьянах. В этих исследованиях были получены однозначные и весьма интересные результаты. Шимпанзе и орангутаны оказались способны к самоузнаванию; остальные приматы (например, все виды нечеловекообразных обезьян) — неспособны, даже проводя сотни часов у зеркала (Gallup, 1994).

рядка; описание становится значительно более дифференцированным и многогранным. Следует отметить, что те же тенденции наблюдаются и в описаниях детьми других людей — к примеру, одноклассников (например, Livesly & Bromley, 1973).

Как указывалось в главе 11, в последних исследованиях когнитивного развития было обнаружено, что многие дошкольники владеют навыками, которые не удавалось выявить в первых исследованиях. Может быть, то же касается и суждений о самом себе? Да, оказывается это так. Четырехлетние дети действительно склонны обращать внимание в основном на внешнее и специфическое, и ни в одном исследовании их самоописания не звучат так же, как самоописания подростков. Однако, если, задавая вопрос, делать акцент на личностных чертах, даже дошкольники иногда описывают себя более психологично и обобщенно. К примеру, есть данные о том, что для того чтобы в описании дошкольника появились указания на характерные черты, достаточно просто переформулировать вопрос (например: «Расскажи, какой ты обычно в школе», а не «Расскажи, какой ты сегодня был в школе») (Eder, 1989). Мы вновь сталкиваемся с существенным обстоятельством: выводы об исследуемых феноменах всегда в той или иной мере зависят от того, как их изучают.

Таблица 12.3

Этапы развития самоописания детей

Возрастной период	Характер описания	Примеры
Раннее детство	Физические характеристики, специфические навыки, то, чем обладают, предпочтения	«У меня рыжие волосы». «Я люблю мороженое»
Среднее детство	Поведенческие особенности, чувства, принадлежность к группе	«Я хорошо отношусь к людям». «Я скаут»
Подростковый возраст	Абстрактные характеристики высокого порядка, личностные особенности	«Я интересуюсь общественной жизнью». «Я хочу нравиться, но боюсь быть отвергнутым»

Самооценка

Методики измерения самооценки в среднем детстве и подростковом возрасте многочисленны и разнообразны (Coopersmith, 1967; Harter, 1985; March, 1988; Piers & Harris, 1969; Rosenberg, 1979). Несмотря на это разнообразие, они имеют ряд общих особенностей. Во-первых, самое очевидное, это то, на что они направлены: нас интересуют оценочный компонент системы Я — какого мнения о себе ребенок, если рассматривать это мнение на оси отрицательного к положительному? Кроме того, для оценки этого мнения обычно используется метод самоотчета — во всех методиках в основе выводов о самооценке лежат в основном ответы детей на прямые вопросы о самоудовлетворенности и самооценности. Третья общая черта — использование множества подобных вопросов: самооценка — глобальный конструкт, и

нам, разумеется, нужно проанализировать как можно больше аспектов жизни ребенка.

Наличие такого количества методик предполагает также и наличие различий между подходами к изучению самооценки. Одно из основных различий имеет отношение к делению на одномерные и многомерные методики. При использовании одномерных методик выявляется единое, общее представление о самооценности — глобальная самооценка, охватывающая множество сторон жизни респондента. Многие из первых методик измерения самооценки по существу — одномерные. Один из известных примеров — опросник самооценки (Coopersmith, 1967, 1981), некоторые вопросы из которого приведены в табл. 12.4. Как можно заметить, использование одномерной методики не означает изучение одного аспекта; напротив, опросник самооценки направлен на выявление мнения испытуемого о себе в разнообразных контекстах и жизненных обстоятельствах. Однако эти контексты или типы даваемых оценок (например, академическая/социальная) не дифференцируются; ответы просто суммируются, и выводится один общий показатель.

Одномерных методик может оказаться достаточно, если люди действительно обладают неким глобальным чувством собственной ценности. Однако, скорее всего, самооценка более дифференцирована. Вполне вероятно, что некий ребенок уверен в себе, когда речь идет о школьных успехах, но недоволен своими отношениями со сверстниками. Именно для выявления подобных потенциальных различий стали разрабатываться многомерные методики.

Таблица 12.4

Примеры вопросов из опросника самооценки Куперсмит

	Похоже на меня	Непохоже на меня
2. Я довольно уверен в себе		
3. Мне часто хочется быть кем-то другим		
17. Я часто сожалею о том, что сделал		
18. Я нравлюсь ребятам моего возраста		
35. Я учусь не так хорошо, как хотел бы		
36. Я способен принять решение и осуществить его		
46. Ребята часто подсмеиваются надо мной		
47. Мои родители меня понимают		

Источник: *The Antecedents of Self-Esteem* (pp. 265–266), S. Coopersmith, 1967, San Francisco: W. H. Freeman and Company.

В табл. 12.5 представлена подборка вопросов одной из наиболее часто используемых многомерных методик измерения самооценки у детей и подростков: Профиля самовосприятия для детей (Harter Self-Perception Profile for Children; Harter, 1985). Как следует из приведенных примеров, с помощью профиля самовосприятия измеряется самооценка в пяти сферах: академическая компетентность, социальное принятие, физическая подготовка, внешний вид и поведение. Каждая из

пяти шкал включает по шесть утверждений, и каждое утверждение предполагает двухфазную модель ответа: ребенок сначала решает, какое из противоположных описаний ему ближе, а затем указывает, насколько точно оно ему соответствует, выбирая квадратик «Немного похоже на меня» или «Очень похоже на меня». Таким образом, каждое утверждение дает определенный показатель по 4-балльной шкале, а сумма из шести показателей отражает самооценку в данной сфере.

Следует ли отказываться от понятия глобальной самооценки с переходом к многомерным методикам? Необязательно. Хартер соглашается с тем, что глобальное ощущение самооценности — валидный и важный конструкт. В профиле самовосприятия глобальный показатель — это не просто сумма ответов по пяти специфическим категориям; глобальное ощущение самооценности составляет отдельную шестую категорию. Последнее утверждение из табл. 12.5 — одно из шести, направленных на непосредственное измерение глобальной самооценки.

Многие интуитивно чувствуют, что самооценка — это нечто многогранное, а не линейное. Однако очевидно, что валидность такого подхода нуждается в эмпирической проверке. Измерение самооценки в разных сферах деятельности имеет смысл только тогда, когда дети, оценивая самих себя, действительно разграничивают эти сферы. Есть данные, свидетельствующие о том, что именно таково положение вещей, и не только при использовании профиля самовосприятия, но и при использовании других многомерных методик (например, March, 1988). Основные данные о валидности этого подхода были получены в корреляционных исследованиях, подобных тем, которые были описаны в главе 4 при рассмотрении понятия конструктивной валидности. Если самооценка действительно многомерна, следует ожидать, что ответы на вопросы, касающиеся одной сферы (например, академической компетентности), будут между собой коррелировать, причем корреляционная связь между ними будет сильнее связи с вопросами, касающимися другой сферы (например, физической подготовки). Именно это явствует из исследований (хотя и не без исключений и спорных моментов — критический анализ методик измерения самооценки см. в Keith & Bracken, 1996). Основной их вывод заключается в том, что самооценка детей (и если уж на то пошло, и взрослых) в одном жизненном аспекте может не совпадать с их самооценкой в другом жизненном аспекте. Высокая или низкая самооценка в отношении одного аспекта развития не гарантирует соответствующей самооценки в других аспектах.

Я уже отмечал, что методики, рассматриваемые в этом разделе, предназначены для изучения самооценки в среднем детстве и подростковом возрасте. Могут ли дети из младшей возрастной группы обладать осмысленной самооценкой? Использование упрощенного варианта «Профиля самовосприятия» говорит о том, что могут. На рис. 12.1 представлен пример из «Картиночной шкалы для оценки воспринимаемой компетентности и социального принятия у маленьких детей» (Pictorial Scale of Perceived Competence and Social Acceptance for Young Children; Harter & Pike, 1984), методики, разработанной для детей 4–7 лет. Как явствует из примера, в «Картиночной шкале» сохраняется двухфазная четырехступенчатая модель ответа «Профиля самовосприятия». В ней также сохраняется разделение на ряд шкал. Представленный пример иллюстрирует утверждение из шкалы когнитивной компетентности; есть также шкалы физической подготовки, материнского принятия и принятия сверстников. Один из альтернативных вариантов от-

Таблица 12.5

Примеры утверждений из «Профиля самовосприятия для детей»

	Очень похоже на меня	Немного похоже на меня				Немного похоже на меня	Очень похоже на меня
1			Некоторые ребята довольны тем, как они учатся	A	Других ребят беспокоит, смогут ли они справиться со школьными заданиями		
2			Некоторым ребятам трудно найти друзей	A	Другим ребятам легко находить друзей		
3			У некоторых ребят получается заниматься всеми видами спорта	A	Другие ребята чувствуют себя на спортивной площадке неуверенно		
4			Некоторые ребята довольны тем, как они выглядят	A	Другим ребятам не нравится, как они выглядят		
5			Некоторым ребятам часто не нравится, как они себя ведут	A	Другим ребятам обычно нравится, как они себя ведут		
6			Некоторые ребята часто недовольны собой	A	Другие ребята вполне довольны собой		

Источник: *Manual for the Self-Perception Profile for Children*, S. Harter, 1985, Denver: University of Denver.

ветов, входящих в последнюю категорию, звучит следующим образом: «У этой девочки довольно много друзей, с которыми она может поиграть. У этой девочки не слишком много друзей, с которыми она может поиграть».

Использование картинок, посредством которых предлагаются альтернативы ответов — одно из явных (и обоснованных) отклонений от строго вербальной процедуры, используемой с более зрелыми респондентами. Еще одно отличие — наличие только четырех, а не пяти специфических шкал, а также отсутствие шкалы глобальной самооценки. Упрощенная форма шкалы в картинках является отражением результатов исследований, в которых изучался вопрос, что различают, а чего не различают маленькие дети, давая оценку самим себе. Даже дошкольники способны рассуждать о себе в оценочных категориях, и даже дошкольники могут неодинаково оценивать себя в разных аспектах своей жизни (например, складывание «пазлов»/игры с друзьями). Однако самооценка маленьких детей все же недостаточно дифференцирована и несколько расплывчата. С возрастом дети становятся все более и более способны разделять и сопоставлять разнообразные аспекты самооценки.

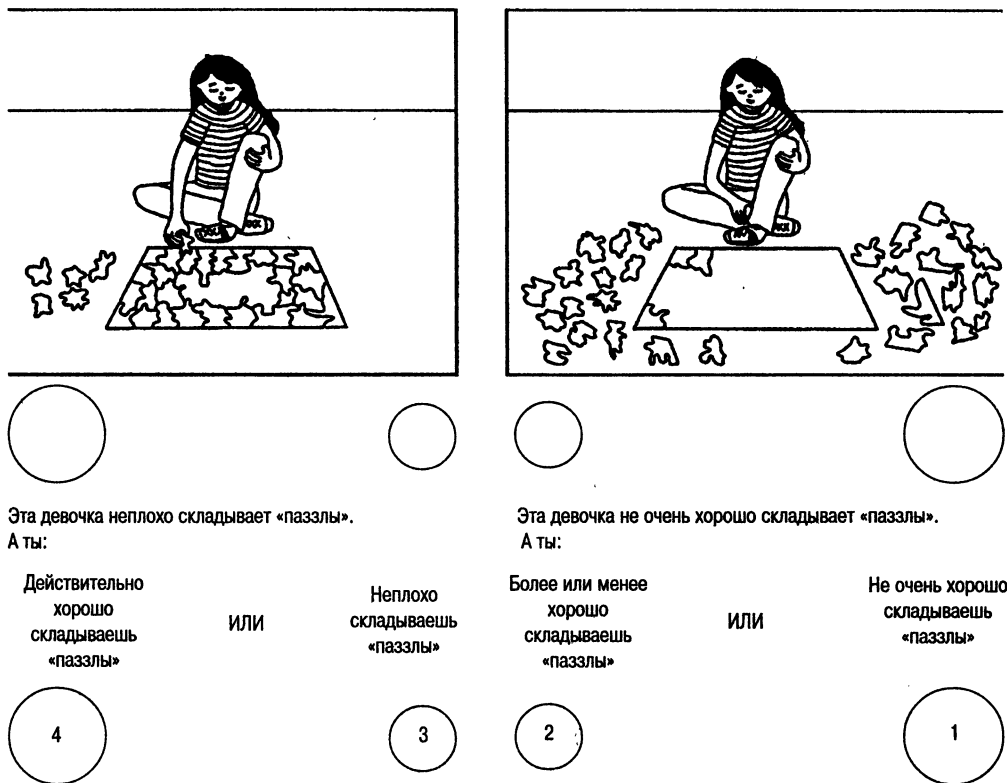


Рис. 12.1. Пример вопроса из «Шкалы в картинках для оценки воспринимаемой компетентности и принятия у маленьких детей». («The Pictorial Scale of Perceived Competence and Acceptance for Young Children», S. Harter & R. G. Pike, 1984. *Child Development*, 55, p. 1973)

Половые различия¹

Общие замечания

Как отмечают Маккоби и Джэклин (Maccoby & Jacklin, 1974), тема половых различий изучается нечасто, в том смысле, что большая часть данных по этому вопросу была получена в исследованиях, в которых не стояла цель анализа этих различий. То есть если сообщаются некие данные о восприятии глубины, сохранении или

¹ В последние годы не раз высказывалось мнение о том, что термины «sex» и «gender» (пол) не следует использовать как взаимозаменяемые понятия; с точки зрения ряда авторов их необходимо употреблять в отношении разных аспектов развития, исторически обозначаемых терминами «половые различия» и «полоролевое развитие». Однако до сих пор по вопросу необходимости сохранения подобного терминологического разделения не существует единого мнения; нет его и по вопросу конкретной разницы между понятиями «sex» и «gender». (Сравните: Deaux, 1993; Gentile, 1993; Unger & Crawford, 1993). В этой главе я придерживаюсь устоявшейся практики многих психологов, использующих эти термины как синонимы.

агрессии, это означает, что исследование проводилось с целью изучения именно этих феноменов. Информацию же о различиях между полами дают буквально все исследования по любой проблематике, если, конечно, выборка (как это бывает в большинстве случаев) включает представителей обоих полов.

Эта специфика данных о половых различиях имеет по крайней мере два следствия. Во-первых, она означает, что база данных для анализа половых различий огромна и неоднородна. Большой объем потенциально значимой информации ставит в чрезвычайно затруднительное положение всякого, кто пытается сделать обзор литературы по этому вопросу и сформулировать на ее основе какие-то выводы. Допустим, что различия между полами анализируются в 200 отчетах исследований феномена сохранения. Из этих 200 исследований, наверное, лишь в дюжине ставилась цель сравнения полов; в других — переменная пола была второстепенной и не имела непосредственного отношения к цели исследования. В одних исследованиях выборки маленькие; в других — большие. В одних исследованиях используются достаточно валидные и надежные методики; методики, применяемые в других исследованиях, вызывают недоверие. Как тогда из этой массы данных выделить главное и отсеять все несущественное? Просто подсчитать количество исследований, в которых были выявлены значимые половые различия и количество исследований, в которых таковых различий не оказалось, а затем сравнить, каких исследований больше, — не самый лучший способ разрешения проблемы. Но любая попытка оценить результаты исследований более дифференцированно требует огромных усилий, профессионализма и тем не менее может привести к результатам, с которыми другие исследователи не согласятся. Яркий пример — критические замечания Блок (Block, 1976) о способе анализа половых различий, который используют Маккоби и Джэклин (Maccoby & Jacklin, 1974).

До этого момента мы говорили в основном о том, как интерпретировать опубликованные отчеты о различиях между полами. Один из важнейших вопросов состоит в том, какая информация вообще имеет больше шансов попасть в печать. Искажения возможны в любом направлении. По мнению Маккоби и Джэклин (Maccoby & Jacklin, 1974), предпочтение отдается информации о наличии различий, а не об отсутствии таковых. Исследователи поэтому склонны отмечать, анализировать и включать в отчеты те случаи, в которых между полами были выявлены различия, и игнорировать те случаи (намного более частые), когда такие различия выявлены не были. Как следствие — преувеличение степени половых различий. Блок же (Block, 1976) утверждает, что искажение может происходить и в обратном направлении. Она отмечает, что многие исследователи рассматривают различие между полами как досадную помеху, которую там, где это возможно, необходимо устранить. Поэтому нередко исследователь предпочитает использовать те методики, применение которых до сих пор не давало оснований для констатации различий между полами. В сомнительных случаях, чтобы удостовериться в том, что ответы представителей обоих полов равноценны, используют пилотажное тестирование. Хотя анализ различий между полами в этом случае становится практически бессмысленным, некоторые журналы и рецензенты требуют включения результатов подобного анализа в отчеты о любых исследованиях, в которых изучалось оба пола. Следствие этого, по мнению Блок, — изобилие бессмысленных отрицательных результатов.

Оба вида искажения, несомненно, существуют; однако никто не знает, насколько они типичны и какое из них имеет большее значение. Главная проблема опять-таки, — второстепенность большей части информации о половых различиях. Этот фактор крайне усложняет любые интерпретации и частично объясняет остроту споров вокруг этой темы.

Обратимся теперь к вопросам, рассматривавшимся в предыдущих главах. В главе 5 мы обсуждали искажающее влияние, которое могут оказать ожидания исследователя на результаты его работы. Как отмечалось, очевидный способ предотвращения подобных искажений — исключение ожиданий, проведение исследования «вслепую», так чтобы тестирующий или наблюдатель не знал о выдвинутых гипотезах и о принадлежности к той или иной группе каждого испытуемого. Со многими переменными проведение исследования «вслепую» вполне осуществимо. Но, как правило, не с полом. Иногда при изучении младенцев от тестирующего или (что более вероятно) наблюдателя удается скрыть пол ребенка. Если ответы для последующего анализа записываются на магнитофон, специалист, производящий оценку, может и не различить голосов мальчиков и девочек. Если ответы перед осуществлением их оценки переводятся в письменный вид, пол испытуемых любого возраста обычно удается скрыть. Однако эти исключения лишь подтверждают правило: обычно наблюдатель, оценивающий поведение, и почти всегда экспериментатор, провоцирующий некое поведение, знают о поле испытуемого.

То, что информация может иметь большое значение, следует из результатов исследований, которые относятся к интересному направлению, известному как исследования типа «Малыш Х». В этих исследованиях изучается эффект экспериментального обозначения ребенка как мальчика или как девочки. Половине взрослых испытуемых сообщают, что малыш, с которым они взаимодействуют или за поведением которого наблюдают, — мальчик; другой половине сообщают, что это девочка. Поскольку во всех случаях ребенок в действительности один и тот же, испытуемые не могут по внешним признакам судить о принадлежности его к определенному полу, и поэтому их ожидания изучаются непосредственно. Такие исследования демонстрируют, что знание о поле может отражаться на поведении взрослого по отношению к ребенку; этот факт непосредственно связан с вопросом об истоках различий между полами и возможности неодинаковых способов социализации (Stern & Karraker, 1989). Для нас же важнее тот факт, что знание о поле влияет также и на интерпретацию взрослым поведения ребенка. В исследовании Кондри и Кондри (Condry & Condry, 1976), к примеру, было обнаружено, что один и тот же поведенческий эпизод (бурная реакция на внезапное появление игрушки из коробки) истолковывался как проявление гнева, если ребенок (по мнению испытуемых) был мальчиком, но как проявление страха, если ребенок был девочкой.

И последнее замечание. Мы обсуждаем эффект знания взрослым пола ребенка. Но верно и обратное: ребенок знает пол взрослого, с которым общается. Иногда дети по-разному ведут себя с мужчиной-экспериментатором и женщиной-экспериментатором. Наиболее проблематична интерпретация половых различий в том случае, когда имеет место не просто главный эффект пола экспериментатора, а взаимодействие фактора пола экспериментатора с фактором пола ребенка. Возможно, к примеру, мальчики лучше реагируют на женщину, проводящую тестирование,

а девочки — на мужчину. Заметьте, что здесь различия между полами есть, однако это различие (как и любое взаимодействие) сложнее, чем кажется на первый взгляд. И, кроме того, оно может быть неправильно истолковано, если тестирование (как это бывает в большинстве случаев) проводят лица одного пола.

Подведем итог сказанному в этом разделе. Определение того, в каких из аспектов психического развития между полами имеются различия, может показаться простой задачей — существует множество психологических методик, и все, что нужно, это применить их к испытуемым обоего пола. Однако, по указанным выше причинам, выявить различия между полами не так просто, поэтому из всех тем психологии развития эта тема вызывает наибольшие разногласия.

Методики изучения половых различий

Главная мысль, высказанная в предыдущем разделе, заключается в том, что для получения информации о половых различиях нет необходимости проводить специальное исследование. Однако, разумеется, есть исследователи, которых в первую очередь интересуют истоки и характер различий между полами. Существует также ряд методик, основная сфера применения которых — исследования полоролевого развития. В этом разделе я приведу лишь несколько из множества примеров подобных методик. (Более подробно см. работы Beere, 1990 и Downs & Langlois, 1988.)

Начнем с исторически самых первых методик, так называемых тестов половой типизации. Голомбок и Фивуш (Golombok & Fivush, 1994, p. 5) определяют половую типизацию как «степень, в которой человек строит свое поведение в соответствии с предписываемыми мужскими и женскими гендерными ролями». Таким образом, точка отсчета — допущение, что гендерные роли существуют, и, как следствие, что существуют и средние различия между полами, то есть особенности мышления, эмоциональной сферы и поведения, которые называют либо «маскулинными» либо «фемининными». Цель тестов половой типизации — определить в течение непродолжительного периода тестирования положение ребенка на шкале «маскулинность—фемининность». Такие тесты могут использоваться для отслеживания изменений уровня половой типизации в период детства. Они могут также использоваться для выявления в половой типизации половых различий как таковых — то есть чтобы выяснить, выражено ли полоролевое поведение мальчика или девочки особенно ярко на определенных этапах развития. Наконец, они могут использоваться для выявления индивидуальных различий в степени половой типизации в рамках одного пола, различий, корни которых можно исследовать в дальнейших исследованиях (например, исследованиях методов родительского воспитания).

Рассмотрим некоторые примеры. Наиболее часто используемые тесты основаны на самоотчетах — то есть мы узнаем о занятиях и предпочтениях детей, напрямую спрашивая их об этом. На рис. 12.2 представлены примеры стимулов из одной подобной методики: «Индекса усвоения половой роли» (*Sex Role Learning Index*, или *SERLI*) (Edelbrock & Sugawara, 1978). *SERLI* предназначен для измерения полотипизированных предпочтений и знаний о полоролевых стереотипах у детей дошкольного возраста. Процедура проводится частично вербально, а частич-

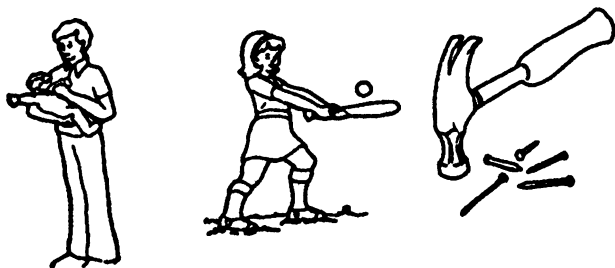


Рис. 12.2. Примеры стимулов из «Индекса усвоения половой роли» (Sex Role Learning Index). («Acquisition of Sex-Typed Preferences in Preschool-Aged Children», С. Edelbrock & A. I. Sugawara, 1978, *Developmental Psychology*, 14, p. 616)

но — с использованием картинок. Для оценки знаний ребенку показывают 10 картинок с изображениями знакомых предметов, половина которых (например, молоток и гвозди, изображенные на рисунке) связана со стереотипически мужскими занятиями, а другая половина (например, иголка и нитки) — со стереотипически женскими занятиями. Во всех случаях вопрос звучит следующим образом: «Кто бы мог использовать [название предмета], чтобы [название деятельности] — мальчики или девочки? Или — и мальчики, и девочки?» Для измерения предпочтений одновременно предъявляют 10 картинок; в этом случае половина картинок иллюстрирует стереотипически мужские занятия (например, игру в бейсбол), а другая половина — стереотипически женские занятия (например, пеленание ребенка). Таких наборов, состоящих из 10 картинок, два; один — для детских занятий, а другой — для ролевого поведения и профессиональных занятий взрослых. В этом случае вопрос ставится следующим образом: «Если бы ты мог делать что-то из этого, что бы ты хотел делать больше всего?» Выбранную ребенком картинку откладывают и вновь повторяют тот же вопрос; таким способом производится полное упорядочение картинок по степени предпочтения. К примеру, если мальчик отдает предпочтение всем пяти «маскулинным» картинкам, игнорируя «фемининные» картинки, он получает максимальный показатель половой типизации.

Как мы уже отмечали, к вербальным отчетам относятся не только самоотчеты; информацию о ребенке можно получить у человека, который его знает. В исследованиях половой типизации такую информацию чаще всего предоставляют родители или учителя. В качестве примера приведем «Опросник занятий дошкольников» (*Pre-School Activities Inventory, PSAI*), разработанную недавно методику, направленную на изучение половой типизации в дошкольном возрасте (Golombok & Rust, 1993a, 1993b). *PSAI* состоит из 24 пунктов, распределенных между тремя сферами: игрушки (например, куклы, игрушечные машинки), занятия (например, игра в дочки-матери, подвижные игры) и личностные характеристики (например: «Нравится красиво одеваться», «Нравятся силовые и подвижные игры»). Как и в *SERLI*, половина утверждений в каждой категории соответствует мужскому стереотипу, а другая половина — женскому. Родитель или учитель оценивает степень соответствия каждого утверждения особенностям ребенка по 4-уровневой шкале, от «никогда» до «очень часто». Таким образом, как и *SERLI*, *PSAI* дает показатель, свидетельствующий о степени соответствия поведения ребенка стереотипу поведения представителей его пола.

Не все методики оценки половой типизации предполагают вербальный отчет: альтернативный вариант — изучать само поведение. Исследования обсервативным методом могут проводиться и в естественных, и специально созданных лабораторных условиях. Исследование, которое мы здесь опишем, проводилось в естественной обстановке. Лэм, Истербрукс и Холден (Lamb, Easterbrooks, & Holden, 1980) в течение 9 недель наблюдали выборку детей в детском саду. Использовалась обсервативная процедура временных срезов, при этом каждого ребенка наблюдали в целом шесть 10-минутных временных периодов. Наблюдение фокусировалось на полотипизированном поведении, и поэтому система кодировки состояла из двух больших категорий: «мужские» занятия и «женские» занятия. Мужские занятия (выделенные на основе проведенных исследований и консенсуса взрослых) включали игру с машинками, лазание, пятнашки и ношение мужской одежды; женские занятия включали игру с игрушечными кухонными принадлежностями, рисование и рукоделие, игру с куклами и ношение женской одежды. Интерес представлял вопрос о том, будут ли по этим категориям мальчики, в среднем по выборке, отличаться от девочек (что в действительности и обнаружилось), и если да, то каковы причины этих столь рано проявляющихся различий. Те результаты исследования, которые имеют отношение ко второму из этих вопросов, мы рассмотрим в разделе «Социализация».

Прежде чем перейти к альтернативным подходам, сделаем несколько общих замечаний, касающихся типов исследований, описанных выше. Во-первых, результат, который получили исследователи (Lamb et al., 1980), имеет более широкое значение: среднестатистические различия между полами по таким параметрам, как предпочтение определенных игрушек и занятий действительно существуют. Именно факт существования подобных различий является основанием для использования этих параметров при оценке степени половой типизации любого ребенка. Во-вторых, по параметру спонтанной игры половые различия проявляются раньше, чем по вербальным параметрам, используемым в таких методиках, как *SERLI*. В некоторых исследованиях игровые предпочтения были обнаружены уже у детей годовалого возраста. В-третьих, полотипизированные предпочтения у мальчиков появляются несколько раньше, чем у девочек; с другой стороны, девочки превосходят мальчиков в отношении знаний о полоролевых стереотипах. Наконец, коэффициент корреляции между показателями ребенка по одной методике измерения половой типизации и показателем по другой методике — весьма умеренный. Таким образом, степень половой типизации поведения нельзя назвать устойчивой.

Одним из поводов для критики описанных выше методик стал тот факт, что в них «маскулинность» и «фемининность» рассматриваются как два полюса линейного измерения. То есть сравнительно высокому показателю по одной совокупности атрибутов, скажем, «маскулинных», неизбежно сопутствует сравнительно низкий показатель по совокупности атрибутов противоположного полюса, в нашем случае, «фемининных» качеств. Однако представляется возможным, что некоторые люди сочетают в себе атрибуты, соответствующие противоположным полюсам традиционной шкалы. Возможно, к примеру, что какой-то человек независим, настойчив (традиционно «мужские» качества) и в то же время отзывчив, заботлив (традиционно «женские» качества).

В последние годы некоторые исследователи пытались преодолеть дихотомию «маскулинность/фемининность», разрабатывая методики, которые допускают андрогинное сочетание черт. Наиболее известная из методик для взрослых — «Полоролевой опросник» (*Bem Sex-Role Inventory*; Bem, 1974) и «Опросник личностных атрибутов» Спенс и Хельмрайх (*Spence and Helmreich Personal Attributes Questionnaire*; Spence & Helmreich, 1978). Обе эти методики, в свою очередь, послужили образцами для детских аналогов. Детская версия методики — «Детский полоролевой опросник» (*Children's Sex Role Inventory*; Boldizar, 1991), а детская версия методики Спенс и Хельмрайх — «Детский опросник личностных атрибутов» (*Children's Personal Attributes Questionnaire*; Hall & Halberstadt, 1980). Ниже основное внимание мы уделим Детскому полоролевому опроснику, или *CSRI*.

Прежде чем рассматривать *CSRI*, необходимо обсудить методику, послужившую его прототипом, — «Полоролевой опросник», или *BSRI*. *BSRI* — методика, основанная на самоотчете и состоящая из 60 описательных утверждений. Двадцать из них отражают качества, которые традиционно считаются (как явствует из результатов стандартизирующей выборки) более типичными и желательными для мужчин (например, самодостаточный, доминантный); 20 — отражают качества, традиционно считающиеся более характерными и желательными для женщин (например, отзывчивый, нежный); и 20 — отражают качества, считающиеся равнотипичными и равножелательными для обоих полов (например, искренний, дружелюбный). Каждое из утверждений испытуемый оценивает по 7-уровневой шкале от «совсем или почти не характерно для меня» до «характерно или почти характерно для меня». Поскольку фемининность и маскулинность оцениваются по отдельности, *BSRI* допускает профиль ответов, который невозможно получить при использовании традиционного биполярного подхода к половой типизации. Особый интерес, как отмечалось выше, представляет так называемый андрогинный профиль, определяемый относительно высокими показателями по шкалам и маскулинности и фемининности.

CSRI имеет много общего с *BSRI*. Он также состоит из 60 пунктов, поровну распределенных между мужской, женской и нейтральной категориями. Утверждения *CSRI* перекликаются с утверждениями *BSRI*, однако формулировка здесь несколько упрощенная, что делает утверждения доступными для понимания детьми (выборку в работе Болдизар составляли дети — учащиеся 3–7 классов) (Boldizar, 1991). Упрощены и шкалы ответов: количество уровней в них было сокращено с 7 до 4. Во вставке 12.7 представлена подборка утверждений из этого теста.

Цель исследования Болдизар (Boldizar, 1991) заключалась в демонстрации валидности *CSRI* как методики измерения у детей уровня маскулинности и фемининности. Поэтому ее отчет служит полезной иллюстрацией сказанному в главе 4 о понятии валидности теста. Болдизар, к примеру, доказала, что ее методика обладает приемлемой надежностью, как ретестовой надежностью, так и внутренней согласованностью. Вспомним, что надежность — необходимое условие валидности. Данные, подтверждающие валидность *CSRI*, как это всегда бывает при установлении конструктивной валидности, разнообразны. Болдизар также доказала, что отдельные шкалы действительно отражают различия между полами: мальчики, в

Вставка 12.7

Примеры утверждений из детского полоролевого опросника**Шкала маскулинности**

Я легко принимаю решения.

Я могу сам позаботиться о себе.

Я имею влияние на многих ребят из моего класса.

Мне нравится делать то, что делают мальчики и мужчины.

Я злюсь, если кто-то встает у меня на пути.

Шкала фемининности

Мне беспокоит то, что происходит в жизни других людей.

Когда кого-то обижают, мне хочется его утешить.

Я отзывчивый человек.

Я добр и заботлив.

Мне нравятся маленькие дети.

Нейтральная шкала

Я честный человек.

Я нравлюсь людям.

Я серьезный человек.

Я стараюсь делать все вовремя.

Я люблю помогать другим людям.

Источник: *Assessing Sex Typing and Androgyny in Children*, J. P. Boldizar, 1991, *Developmental Psychology*, 27, p. 509.

среднем, имеют более высокий показатель по шкале маскулинности, а девочки — по шкале фемининности. Она также продемонстрировала наличие корреляционной связи между показателями маскулинности/фемининности и ожидаемым поведением при использовании традиционной методики предпочтения определенных игрушек и занятий. Корреляция была обнаружена также с описанным ранее «Профилем самовосприятия» Хартер; был проверен и подтвержден ряд связей между самооценкой и половой типизацией, прогнозировавшихся на основе теоретических положений и эмпирических исследований. Помимо индексов конвергентной валидности посредством демонстрирования того, что показатели CSRI не связаны с IQ и академическими достижениями, была подтверждена дивергентная валидность. Наконец, на основе данных выборки взрослых был получен коэффициент корреляции между CSRI и BSRI, равный 0,88, — что и следовало ожидать, учитывая, что обе эти методики измеряют один и тот же конструкт.

Условно говоря, между последним типом методик, которые мы рассмотрим, можно провести аналогию с подходами, обсуждавшимися в последней части раздела «Нравственное развитие». Там нас интересовала когнитивная сторона нравственности: методы изучения того, как дети рассуждают о вопросах нравственности. Здесь в центре внимания — когнитивная сторона половой типизации: как ребенок рассуждает о половых ролях и половых различиях, а также, как его рассуждения изменяются в процессе развития.

Какого рода вопросы рассматриваются под рубрикой «когнитивный аспект половой типизации»? Простейший из этих вопросов касается способности ребенка правильно определять пол. На каком этапе развития мальчик начинает называть себя мальчиком, а девочка — девочкой? Когда ребенок начинает определять пол других людей? Знает ли 3-летний ребенок, кто из его знакомых сверстников мальчик, а кто — девочка? Допустим, различие между людьми, пол которых нужно определить, и ребенком более значительное: это взрослые, а не дети. Осознает ли маленький ребенок, что папа это «мальчик», а мама это «девочка»?

Разнообразные данные свидетельствуют о том, что способность правильно определять пол формируется достаточно рано (Fagot & Leinbach, 1993). Обозначение пола — типичный элемент лексикона большинства маленьких детей, а обозначение собственного пола (например: «Я мальчик»), как правило, — одно из первых проявлений формирующейся Я-концепции ребенка. К 2,5–3-летнему возрасту практически все дети в состоянии правильно ответить на прямой вопрос: «Ты девочка или мальчик?» (Slaby & Frey, 1975). К этому возрасту дети также начинают правильно называть пол других людей, отвечая, к примеру, на вопрос о том, кто изображен на картинке, мальчик или девочка, мужчина или женщина (Leinbach & Fagot, 1986). Однако интересно отметить, что ребенку назвать пол взрослого обычно легче, чем назвать пол ребенка, предположительно, по причине большей выраженности у взрослых отличительных особенностей пола (например, черт лица, причёски). Стоит также отметить, что существуют заметные индивидуальные различия в темпах усвоения навыков определения пола: возраст, в котором ребенок начинает справляться с заданиями на определение пола, колеблется от 24 до 40 месяцев (Leinbach & Fagot, 1986).

Узнав, что ребенок правильно определяет пол, мы, естественно, сталкиваемся со следующим вопросом: какие критерии использует ребенок при определении пола? У взрослых, разумеется, такими критериями являются анатомические отличия, хотя в большинстве случаев бывает достаточно других, более очевидных признаков (например, длина волос, тембр голоса, одежда). Результаты исследований свидетельствуют о том, что ребенок сначала ориентируется на признаки, наиболее доступные для восприятия. Томпсон и Бентлер (Thompson & Bentler, 1971) оценивали значение, которое при определении пола придавали разным признакам испытуемые четырех возрастов: 4-летние, 5-летние, 6-летние и взрослые. Стимулами служили раздетые куклы, различавшиеся по трем параметрам: тип гениталий (мужские или женские), телосложение (мужское телосложение, описанное как «мускулистое, крепкое» или женское телосложение, описанное как «тело с развитыми грудью и бедрами») и длина волос (длинные или короткие). Использовались все возможные комбинации значений каждого из трех параметров — то есть 8 кукол. Это, естественно, означало, что у некоторых кукол все параметры гармонизировали (например, мужские половые органы, мужское телосложение и короткие волосы), тогда как у других — не соответствовали друг другу (например, мужские половые органы, женское телосложение и длинные волосы). Каждому ребенку показывали по одной подобной кукле и просили определить ее пол. Томпсон и Бентлер обнаружили, что в сомнительных случаях только взрослые рассматривали половые органы как решающий фактор. Хотя дети в той или иной мере учиты-

вали все признаки, основное, на что они ориентировались, были волосы, за которыми следовало телосложение. То есть большинство детей считало куклу с мужскими половыми органами, но женским телом и длинными волосами куклой женского пола; и наоборот, пол куклы с женскими половыми органами, но короткими волосами довольно часто определялся как мужской.

Подобные исследования свидетельствуют о том, что когнитивные навыки определения пола у маленьких детей еще недостаточно развиты. Этот вывод становится еще более обоснованным, если рассмотреть результаты исследований, направленных на изучение других аспектов понимания пола: осознание того, что пол — неизменное качество. В действительности, именно некоторые положения теории Колберга (Kohlberg, 1966), касающиеся понятия о постоянстве пола, вызвали интерес к когнитивной стороне половой типизации. В своей работе 1966 года Колберг выдвинул две гипотезы. Первая — дети лишь постепенно начинают осознавать, что пол — это неизменное качество, на котором не отражаются возрастные изменения, желания или конкретные обстоятельства (например, ношение определенной одежды или определенной прически). Колберг уподобил этот процесс постепенного осознания инвариантности пола развитию представлений о константности физического мира, в частности, известному пиажетианскому понятию о сохранении. Вторая гипотеза Колберга заключалась в том, что в развитии половой типизации ведущую роль играет осознание ребенком константности пола. Мальчик начинает идентифицировать себя с отцом и усваивать мужские предпочтения и атрибуты, только осознав, что он мужского пола и его пол всегда останется мужским.

С момента опубликования этой работы Колберга было проведено множество десятков исследований, направленных на изучение возрастных изменений в понимании константности пола. Примером служит исследование Маркус и Овертон (Marcus & Overton, 1978). На рис. 12.3 изображены два стимула из стимульного материала, использовавшегося в этом исследовании (этот стимульный материал был разработан Emmerich & Goldman, 1972). Листы со схематическими изображениями фигур были сброшюрованы и разрезаны так, что, переворачивая их, можно было совместить голову мальчика с фигурой девочки и голову девочки с фигурой мальчика. Каждому испытуемому (детям 5–7 лет) задавалось по пять вопросов. Начиная с картинки, соответствующей полу ребенка, экспериментатор последовательно производил трансформации с прическами фигур, их одеждой и одновременно — прической и одеждой. Вслед за каждым изменением ребенка спрашивали, осталась ли фигура девочкой (или мальчиком) или пол изменился. Ребенок также отвечал на вопрос о том, изменился ли бы у фигуры пол, если бы ей стали нравиться другие игрушки или она «сильно захотела» бы стать другого пола.

Описанная выше процедура — один из методов изучения представлений о константности пола. Она имеет заметное сходство с заданиями Пиаже на сохранение: в обоих случаях производятся трансформации внешнего облика, которые могут ввести ребенка в заблуждение; и в обоих случаях ребенок должен абстрагироваться от непосредственно данной видимости, с тем чтобы осознать неизменность сути. В другом методе более ярко выражен вербальный компонент. В табл. 12.6 представлены примеры вопросов из протокола интервью, разработанного Слэби и Фрей (Slaby & Frey, 1975) и часто используемого при изучении представлений о кон-

стантности пола. В процедуре этих исследователей не ставится задача абстрагирования от обманчивого внешнего вида; изменения остаются гипотетическими и описываются в серии вопросов, направленных на проверку знаний ребенка о константности пола в разных ситуациях.

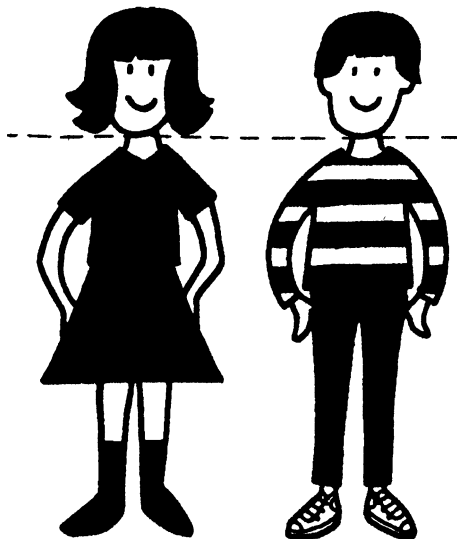


Рис. 12.3. Схематические изображения, используемые для оценки представлений ребенка о константности пола. («Boy-Girl Identity Task», W. Emmerich and K. S. Goldman, 1972. In V. Shipman (Ed.), *Disadvantaged Children and Their First School Experiences* (ETS PR 72-20), Princeton, NJ: Educational Testing Service)

Таблица 12.6

Примеры вопросов из теста представлений о поле Слэби и Фрей

Когда ты был совсем маленьким, ты был мальчиком или девочкой?

Ты когда-нибудь был маленьким/маленькой [пол, противоположный указанному при ответе на первый вопрос]?

Когда ты вырастешь, то будешь мамой или папой?

Мог бы/могла бы ты когда-нибудь быть [пол, противоположный указанному при ответе на первый вопрос]?

Если бы ты носил одежду для [пол, противоположный полу ребенка, то есть «для мальчиков» или «для девочек»], ты бы был девочкой или мальчиком?

Если бы ты носил одежду для [пол, противоположный полу ребенка], был бы ты [пол, противоположный указанному при ответе на первый вопрос]?

Если бы ты играл в игры для [пол, противоположный полу ребенка], ты бы был девочкой или мальчиком?

Если бы ты играл в игры для [пол, противоположный полу ребенка], был бы ты [пол, противоположный указанному при ответе на первый вопрос]?

Мог бы ты быть [пол, противоположный полу ребенка], если бы этого захотел?

Источник: *Development of Gender Constancy and Selective Attention to Same-Sex Models*, R. G. Slaby and K. D. Frey, 1975, *Child development*, 46, p. 851.

Кратко рассмотрим несколько результатов, полученных при использовании таких методик. Во-первых, Колберг оказался прав в том, что понятие о константности пола осваивается в ходе развития: маленькие дети обычно не справляются с заданиями на константность пола, а способность к их выполнению появляется в возрасте около 4–7 лет. Во-вторых, из того, что этот возрастной диапазон довольно широк, можно сделать вывод о значении метода оценки представлений о константности пола. Как и следовало ожидать, дети обычно лучше справляются с вербальной процедурой, такой как задания Слэби и Фрей, а не с заданиями, в которых они сталкиваются с обманчивой видимостью несоответствия признаков пола, как в методике Маркус и Овертон. Форма константности также имеет значение. Большинство детей начинают осознавать постоянство пола во времени несколько раньше, чем его неизменность в разных ситуациях. Первые два вопроса из табл. 12.6 направлены на оценку представлений о стабильности пола, а следующие два — на оценку представлений о постоянстве пола. Последний вывод, связанный со вторым теоретическим предположением Колберга: между пониманием постоянства пола и полотиципизированным поведением существует лишь умеренная и неустойчивая связь (Beal, 1994). Таким образом, оказывается, что понимание константности пола в какой-то мере определяет полоролевое развитие, но далеко не полностью.

Как отмечалось, теория Колберга (Kohlberg, 1966) о константности пола послужила стимулом к изучению когнитивного аспекта половой типизации. В последние годы когнитивный подход вышел за рамки изучения константности пола. Сегодня особый интерес вызывает понятие **гендерной схемы** (Martin, 1993). Схема — это мысленная репрезентация некоего класса событий, то есть гендерная схема — это упорядоченные знания и убеждения ребенка, имеющие отношение к полу. Как следует из этого определения, гендерная схема — это широкое понятие, охватывающее разнообразные специфические представления. Она включает навыки определения пола, а также представления — сначала ошибочные, а затем правильные — о константности пола. Она включает также знания о полоролевых стереотипах, для оценки которых предназначен SERLI и похожие методики, — какие игрушки предпочитает каждый пол, какие личностные качества характерны для мальчиков, а какие для девочек? В целом, гендерная схема включает представления о самом себе и о других, о собственной половой принадлежности и о половой принадлежности других, о детях и о взрослых. То есть это значительно более сложное понятие, чем константность пола, и развитие отдельных аспектов гендерных схем начинается значительно раньше, чем развитие представлений о константности пола.

Возможно, вы заметили, что понятие схемы и понятие сценария, рассмотренное в главе 11, имеют много общего. В действительности, сценарий — это частный случай схемы, имеющий отношение к структуре знакомых событий. В главе 11 мы приводили вывод, полученный на основе исследований сценариев и заключающийся в том, что, сформировавшись, сценарий начинает влиять на процесс восприятия, интерпретации, запоминания и сохранения новой информации. То же касается и схем в целом, и гендерных схем в частности. Было, к примеру, обнаружено, что дети лучше запоминают те события, которые согласуются с их гендерными схемами, и в некоторых случаях даже искажают информацию, приводя ее в соответствие с собственными ожиданиями, — сообщая, к примеру, что врач был мужчиной, а медсестра — женщиной, когда в действительности все было наоборот (Carter &

Levy, 1988; Liben & Signorella, 1993). Гендерные схемы влияют и на поведение. Например, игры детей, которые относительно рано начинают употреблять в своей речи обозначения пола, более полотиципизированны, чем игры детей, не знающих слов, обозначающих принадлежность к определенному полу (Fagot, 1985). То есть в целом второе положение когнитивно-структуралистской теории Колберга (Kohlberg, 1966) подтверждается: то, что дети думают о поле, влияет на то, как они себя ведут, а возрастные изменения в их представлениях в значительной мере определяют изменения в их поведении.

Социализация

Обратимся теперь ко второму общему вопросу, который мы должны рассмотреть в этой главе: к вопросу о детерминантах. Чем определяется социальное поведение ребенка: и те его формы, которые мы уже обсуждали, и другие? Излагая этот материал, основное внимание я уделю процессам социализации — тому, как другие люди (родители, учителя, братья и сестры, сверстники) влияют на социальное развитие ребенка. Будут рассмотрены три подхода к изучению социализации: лабораторные исследования, обсервативные методики и методики вербального отчета или рейтинг. Эта структура изложения, как вы помните, использовалась и в первой части данной главы.

Прежде чем начать, сделаем небольшую оговорку. Хотя опыт социализации, который мы рассмотрим, весьма существенен, его недостаточно, чтобы объяснить все особенности социального развития ребенка. Несомненно, определенную роль играют биологические факторы. Полоролевые различия частично (но именно частично!) обусловлены биологически. То же касается привязанности и темперамента, двух важнейших социальных феноменов, о которых шла речь в главе 10. Определенную роль играет когнитивный компонент. Как следует из материала, посвященного нравственности и половой типизации, когнитивные факторы приобретают все больший вес при объяснении феноменов социального развития. Освоение ребенком определенных когнитивных навыков лишь частично объясняется рассматриваемыми здесь социальными процессами. Иными словами, изложение опять-таки будет избирательным: в этом разделе обсуждаются лишь некоторые факторы социального развития.

Лабораторные исследования

Цель лабораторных исследований социализации — создание экспериментального аналога жизненного опыта социализации. Конечно, в лаборатории невозможно воспроизвести естественные условия во всей их сложности. Однако можно выделить некие ключевые процессы — воспроизвести определенные ситуации с той мерой точности, которая бы позволяла генерализовать полученные результаты. При этом контроль и точность лабораторного исследования дают возможность выявить причинно-следственные связи.

Рассмотрим два примера. Даже без проведения формального исследования ясно, что одним из важнейших факторов социального развития являются разнообразные модели, с которыми сталкивается ребенок. Дети усваивают разные формы

поведения, наблюдая их у других людей. Наблюдая за поведением других людей, ребенок также может изменить уже существующие у него модели поведения. В современных работах по социальному научению особенно подчеркивается роль научения через наблюдение (Bandura, 1989).

Как его изучать в лабораторных условиях? Возьмем в качестве примера агрессию. Агрессия, несомненно, самая популярная тема среди занимающихся экспериментальным исследованием моделирования; поэтому недостатка в исследованиях, которые можно было бы привести в качестве примера, нет. Остановимся на одном из таких экспериментов (Hicks, 1968). В исследовании Хикс женщина-экспериментатор по одному приглашала в игровую комнату испытуемых (детей 5–12 лет). По пути в игровую комнату экспериментатор вспоминала, что забыла кое-что сделать в библиотеке, и предлагала ребенку, пока она отсутствует, посмотреть короткий фильм. В фильме, который демонстрировался на киноэкране в помещении, находившемся рядом с библиотекой, показывался взрослый мужчина, совершавший различные агрессивные действия по отношению к кукле Бобо. По окончании фильма ребенка отводили в игровую комнату, где он играл в течение 15 минут. В комнате находились игрушки, которые можно было использовать для подражательной или неподражательной агрессии (включая куклу Бобо), а также игрушки, используемые в нейтральных играх. За поведением ребенка наблюдали через одностороннее зеркало и регистрировали все (довольно частые) случаи агрессии.

Эксперимент Хикс с предъявлением образца поведения типичен в нескольких отношениях. Во-первых, показ образца для подражания связан с тем, что называют «экспозиционной техникой». Ребенка не просят напрямую следить за поведением модели; он просто присутствует в тот момент, когда имеет место определенное поведение. Аналогично, тест с подражанием связан с предоставлением возможности подражать, а не прямым инструктированием; ребенок сам решает, как себя вести. Можно утверждать, что такая процедура соответствует большинству реальных жизненных ситуаций, которые нас интересуют; ситуаций, когда ребенок наблюдает поведение других, но подражать ему его никто не заставляет. Наконец, этот эксперимент типичен еще и в том смысле, что в нем допускается наличие и подражательных, и неподражательных действий. Дети били куклу Бобо молотком так же, как это делал взрослый; но они также демонстрировали ряд агрессивных действий, которые модель не совершала. Таким образом, исследование демонстрирует два следствия влияния образца на поведение: точное воспроизведение поведения и связанные с образцом действия, уже имеющиеся в репертуаре ребенка.

Как подчеркивалось ранее, одно из преимуществ лабораторного подхода заключается в том, что он допускает проведение систематических манипуляций с разнообразными потенциально значимыми независимыми переменными. В исследованиях с научением через наблюдение изучалось множество таких переменных. Нас, к примеру, могут заинтересовать характеристики лица — образца для подражания, которые влияют на поведение ребенка. В исследовании Хикс моделью был взрослый мужчина; в других исследованиях сопоставляется воздействие модели-женщины и модели-мужчины (например, Perry & Perry, 1975) или модели-взрослого и модели-ребенка (например, Bandura, Grusec, & Menolve, 1967). В исследовании Хикс модель появлялась на экране; во многих исследованиях дети видят живую

модель, а в некоторых исследованиях сравнивается воздействие образцов поведения на телеэкране и образцов поведения, демонстрируемых непосредственно (например, McCall, Parke, & Kavanaugh, 1977). Манипуляции возможны и с отношениями между ребенком и лицом, служащим образцом поведения. В одних исследованиях, в том числе и в исследовании Хикс, модель — незнакомый ребенку человек. В других исследованиях изучается влияние разного рода взаимоотношений между лицом, демонстрирующим образец поведения, и ребенком (например, теплые, основанные на поддержке / холодные, отчужденные) либо посредством привлечения человека, известного ребенку (например, Hetherington & Frankie, 1967), или посредством предварительного взаимодействия между моделью и ребенком (например, Jeffrey, Hartmann, & Gelfand, 1972). Интерес представляет также роль событий, следующих за демонстрированием определенного образца поведения. Можно, к примеру, попытаться выяснить, наблюдается ли у детей склонность подражать поведению, которое, как они увидели, подкрепляется, и избегать поведения, которое, как они видели, наказывается (например, Bandura, Ross, & Ross, 1963b; ответ на этот вопрос, кстати, утвердительный). Наконец, экспериментальные манипуляции могут производиться с разнообразными ситуационными аспектами и фазой предъявления образца поведения/фазой подражания. В исследовании Хикс, к примеру, было включено экспериментальное условие, при котором вместе с ребенком агрессивный фильм смотрел взрослый мужчина, который комментировал происходящее на экране. У половины детей взрослый присутствовал в фазе подражания, а у другой половины — нет.

Пока мы обсуждали разнообразные манипуляции, производимые в исследованиях моделирования и направленные только независимые переменные. Метод моделирования также позволяет изучать эффекты или зависимые переменные, помещенные в разного рода контексты. В лабораторных исследованиях моделирования было продемонстрировано влияние модели на десятки форм поведения. По количеству исследований вслед за агрессией, пожалуй, идут разнообразные формы так называемого нравственного поведения. Эффект моделирования был продемонстрирован на всех поведенческих проявлениях, описанных в первой части этой главы — мошенничество, способность устоять перед соблазном, склонность делиться, оказывать помощь и т. д. Более того, влияние обнаружилось как в «желательном», так и в «нежелательном» направлениях. Ребенок, наблюдавший, как кто-то делится, к примеру, в дальнейшем с большей готовностью поделится сам; ребенок, наблюдавший, как модель отказывается поделиться, впоследствии также делится неохотно. Аналогично, ребенок, наблюдавший, как кому-то удается устоять перед соблазном, сам приобретает большую устойчивость; а ребенок, наблюдавший, как кто-то уступает соблазну, сам впоследствии склонен уступать (обзор литературы по этому вопросу см у Perry & Bussey, 1984).

Подведем краткие итоги достоинств лабораторного подхода к изучению моделирования. В лабораторных исследованиях анализируется процесс, который играет существенную роль в социализации ребенка: влияние поведения других людей на поведение ребенка. Осуществление контроля — контроля не только формы моделирования, но и других аспектов ситуации — позволяет выявить в лабораторных условиях причинно-следственные связи между моделированием и последующим

поведением. Наличие контроля, кроме того, делает возможным рассмотрение широкого диапазона независимых переменных.

Учитывая то, о чем говорилось в главе 6 и в первой части этой главы, мы должны остановиться и на недостатках лабораторных исследований. Как всегда, наибольшее беспокойство вызывает внешняя валидность. Что демонстрирует большинство лабораторных исследований научения через наблюдение, так это то, что в строго заданных условиях поведение человека может влиять на поведение детей в не менее строго заданных условиях сразу после фазы наблюдения. В естественной среде интерес представляют стойкие, отдаленные последствия моделирования; по результатам большинства лабораторных исследований мы можем лишь догадываться о подобных отсроченных последствиях. В естественной среде воздействие образцов поведения чаще всего сочетается с воздействием множества других стимулов, поэтому ребенок должен каким-то образом заметить ее и перенести эффекты ее влияния на другую, возможно совершенно иную ситуацию. В большинстве лабораторных исследований и фаза наблюдения, и фаза реагирования крайне упрощены, то есть упускаются из вида факторы, которые имеют место в естественной среде, такие как внимание, избирательность и перенос. Наконец, в большинстве лабораторных исследований роль модели выполняет незнакомый или почти незнакомый ребенку человек. До какой степени возможен перенос результатов таких исследований в домашнюю обстановку, где ребенок бесчисленное количество раз в течение многих лет наблюдает поведение людей, с которыми он находится в близких эмоциональных отношениях, — спорный вопрос.

Второй процесс социализации, который мы рассмотрим — это обоснование предъявляемых требований. Идея о том, что обоснование взрослым своих требований к ребенку является эффективным методом дисциплинирования, нашла свое выражение во многих теориях социализации и согласуется с взглядами многих родителей. Это то, что, по нашему мнению, должно было бы соответствовать действительности: обоснование требований должно повышать вероятность следования ребенком этим требованиям. Но подтверждают ли эмпирические исследования это распространенное мнение? Да, исследования неизменно его подтверждают; в действительности, результаты многочисленных исследований свидетельствуют об одном и том же: обоснование требований, предъявляемых ребенку, оказывает благотворное действие (Grusec & Lytton, 1988). Здесь я сосредоточусь на одном из источников данных: на лабораторных исследованиях переменной обоснования предъявляемых требований.

В начале этой главы была описана типичная процедура оценки способности ребенка не поддаваться соблазну нарушить запрет: тест «Запрещенные игрушки». В своей базовой форме — которая была представлена — тест «Запрещенные игрушки» довольно прост: взрослый налагает запрет, а вопрос состоит в том, будет ли ребенок в состоянии его соблюдать. Вас, наверное, не удивит, что многим детям не удавалось успешно пройти это испытание: если не считать авторитета взрослого, у ребенка нет оснований отказывать себе в игре. Исследования фактора обоснования направлены на прояснения вопроса о том, что происходит, если ребенку объяснить причину запрета. Послушнее ли станут дети, если запрет будет сопровождаться обоснованием — объяснением, почему нельзя трогать игрушки? Как уже указывалось, ответ на этот вопрос утвердительный: результаты исследований свидетельствуют о

том, что обоснование повышает действенность запрета (Parke, 1981). Было продемонстрировано, что соблюдению ребенком запрета взрослого способствует, к примеру, указание на то, что игрушки могут сломаться, или обращение внимания на то, что игрушки принадлежат другому ребенку (Parke & Murray, 1971). Послушания можно добиться указанием на то, что экспериментатор будет недоволен, если ребенок ослушается, а также указанием на то, что экспериментатора расстроит или разочарует неспособность ребенка отказать себе в игре (Kuczynski, 1983; Sawin & Parke, 1980).

Как и в случае с научением через наблюдение, наглядное доказательство того, что обоснование может оказаться эффективным, — лишь первый шаг исследовательской программы. Вновь мы можем исследовать вариации базового метода, изменяя параметры, связанные либо с независимой, либо с зависимой переменной. Как следует из приведенных выше примеров, один из важнейших вопросов касается относительной эффективности разных типов обоснований — к примеру, объектно-ориентированного/личностно-ориентированного. Еще один вопрос касается возрастных изменений: повышается ли эффективность обоснования в процессе развития ребенка? Ответ в целом утвердительный, однако переменная возраста взаимодействует с переменной типа обоснования. Для дошкольников лучше всего подходят простые обоснования (например: «Игрушка может сломаться»), с возрастом эффективнее становятся более сложные объяснения (например, указание на право собственности, на необходимость сопереживания). Обоснование можно использовать в сочетании с другими методами дисциплинирования — к примеру, наказания ребенка за то, что он нарушил уговор не трогать игрушки (в этом исследовании наказание обычно принимает форму громкого звонка, иногда в сочетании с нестрогим вербальным порицанием). Обоснование действительно повышает эффективность наказания; в действительности, само по себе обоснование требований обычно эффективнее наказания. Наконец, влияние обоснования требований можно изучать не только на примере способности устоять перед соблазном, но и изучая разнообразные формы просоциального поведения, к примеру, склонность делиться (Perry, Bussey, & Freiberg, 1981) или предложение помощи нуждающимся в поддержке (Staub, 1971). Как показывают эти исследования, обоснование не всегда эффективно (в действительности, мало что эффективно в этом случае); в целом, однако, обоснование требований стимулирует поведение, которое хочет вызвать агент социализации.

Натуралистическое наблюдение

В первой части главы рассматривались исследования таких важнейших аспектов социализации, как нравственное поведение и половая типизация. В исследованиях, которые мы обсудим здесь, изучаются в том числе и описанные выше феномены. Однако в них, кроме того, осуществляется попытка оценить естественный опыт социализации. Цель состоит в установлении причинно-следственной связи между определенным опытом и определенными психическими явлениями.

Приведем два примера. В предыдущем разделе основное внимание было уделено роли взрослого как образца поведения и влиянию обоснования выдвигаемых им требований. Однако взрослые — не единственные агенты социализации, с которы-

ми сталкиваются дети; на ребенка могут влиять и другие дети. А научение через наблюдение и обоснование требований — далеко не единственные механизмы социализации ребенка. Не нужно быть приверженцем теории Скиннера, чтобы догадаться, что третьим важнейшим механизмом социализации ребенка является подкрепление поведения ребенка. Первый пример в связи с этим будет иметь отношение к такой переменной, как подкрепление со стороны сверстников.

Сверстники могут подкреплять и подкрепляют разнообразные формы поведения. Майкл Лэм и его коллеги (Lamb et al., 1980; Lamb & Roopnarine, 1979) исследовали роль подкрепления со стороны сверстников в формировании поведения, соответствующего стереотипам поведения для обоих полов. Один из компонентов данного исследования уже был описан в этой главе: обсервативная оценка полотицизированного поведения у детей дошкольного возраста. Использовался и еще один набор обсервативных методик, в которых оценивалась реакция детей на полотицизированное поведение их сверстников. Как реагируют другие дети на упомянутые ранее маскулинные и фемининные формы поведения ребенка, и изменяется ли их реакция как функция от соответствия этой формы поведения полу ребенка (например, кто произвел «маскулинное» действие, мальчик или девочка)? Интерес представляет как подкрепление, так и наказание. Было выделено семь категорий реакции, которые рассматривались как подкрепление; среди них: «похвала», «одобрение», «согласие» и т. д. Оценивалось также пять типов наказания, в том числе «критика», «неодобрение».

Было сделано несколько выводов. Реакция сверстников определенно зависела от соответствия поведения ребенка его полу. И мальчики, и девочки чаще всего получали подкрепление со стороны детей, если их поведение соответствовало типичной для их пола модели поведения. Кроме того, другие исследователи (например, Fagot, 1977) наблюдали подобные дифференцированные реакции на полотицизированное поведение среди детей начальных классов. Майкл Лэм с соавторами приводят также данные, свидетельствующие о том, что реакции, квалифицировавшиеся как подкрепление и наказание, действительно влияли на дальнейшее поведение ребенка. На прерывание подкреплявшегося действия в среднем требовалось значительно больше времени, чем на прерывание действия, результатом которого было наказание; и в целом, продолжительность эпизодов поведения, получавшего подкрепление, была больше, чем продолжительность эпизодов поведения, получавшего наказание или игнорировавшегося. Майкл Лэм с соавторами обнаружили также, что результат подкрепления или наказания частично зависел от соответствия поведения ребенка модели поведения, типичной для его пола. Подкрепление было эффективнее, если поведение соответствовало этой модели; аналогично, наказание действовало быстрее, если поведение этой модели не соответствовало. Майкл Лэм с соавторами предлагают когнитивную интерпретацию этого факта. Они полагают, что уже 3-летние дети осознают полоролевые стандарты и руководствуются ими; таким образом, подкрепление или наказание со стороны сверстников служит преимущественно напоминанием об этих стандартах.

Для рассмотрения следующего примера натуралистического исследования социализации переключим внимание с отношений между детьми в детском саду на взаимоотношения ребенка с родителями и сиблингами. В серии статей Паттерсон

(Patterson, 1979; Patterson & Bank, 1989; Patterson, Reid, & Dishion, 1992) опубликовал результаты долгосрочного исследования детерминант агрессивного поведения детей в домашней обстановке. Первичной выборкой для исследования послужили семьи, которым было рекомендовано участие в орегонском «Проекте социального научения» (Oregon «Social Learning Project») в связи с тем, что, по крайней мере, один из детей в этих семьях демонстрировал высокий уровень агрессии. Цель исследования состояла в выявлении процессов, обуславливающих агрессивное поведение и имеющих отношение не только к самому ребенку, вызвавшему беспокойство, но и ко всем членам семьи. Исследование проводилось методом натуралистического наблюдения — многочасового наблюдения в естественной домашней обстановке за агрессивным поведением, за предшествующими ему событиями и за его последствиями.

Остановимся на одном из отчетов по этому проекту (Patterson & Cobb, 1971). Испытуемыми в этой части исследования были 24 семьи, в каждой из которых был, по крайней мере, один гиперагрессивный мальчик в возрасте от 6 до 13 лет. В целом каждую семью наблюдали на протяжении 10 часов — 5 дней в неделю, по 1 часу в день в течение 2 недель. Регистрировалось поведение каждого члена семьи и реакция на это поведение других членов семьи. Система кодировки занимала промежуточное положение между макро- и микроуровнями, рассмотренными ранее. В нее входило 30 категорий, по мысли авторов отражающих разнообразные аспекты социального взаимодействия: близость, внимание, гиперактивность, деструктивность, приказ, зависимость, запрет, игнорирование, игра, крик, негативизм, неодобрение, несогласие, нормативное поведение, одобрение, отсутствие реакции, плач, поддразнивание, позитивное физическое воздействие, прикосновение, принятие, работа, разговор, самостимуляция, смех, снисхождение, согласие, удар, унижение и хныканье. Паттерсон и Кобб (Patterson & Cobb, 1971) больше всего интересовала категория Удар, а также факторы-побудители и подкрепление ударов.

Процедура анализа, которую использовали Паттерсон и Кобб, включала подсчет условной вероятности. Первичными данными для ее подсчета служит исходный уровень неких действий, в данном случае ударов. Исходный уровень отражает частоту неких действий, то есть общее количество всех зафиксированных поведенческих проявлений, разделенное на количество поведенческих актов определенного типа. Если, к примеру, было зафиксировано 10 000 поведенческих проявлений и 80 из них относились к категории Удар, исходный уровень для категории Удар будет 0,008. Установив исходный уровень, можно анализировать изменения вероятности поведенческого проявления как функцию от воздействия предшествующих этому поведенческому проявлению стимулов, а изменение вероятности того, что текущее поведение будет продолжаться или повторится, можно проанализировать как функцию от разнообразных последствий. Сосредоточимся здесь на событиях, предшествующих нанесению ударов. Паттерсон и Кобб проанализировали вероятность нанесения ударов как функцию от присутствия каждого из членов семьи. Не вызывает удивления тот факт, что, как обнаружилось, присутствие матери или отца было связано со значительным снижением вероятности нанесения ударов относительно исходного уровня — 6-кратное снижение в случае присутствия отца. И напротив, присутствие младшей сестры или старшего брата было

связано со значительным повышением вероятности нанесения ударов. Предположительно, эти эффекты обусловлены тем, что родители и сиблинги в неодинаковой мере провоцируют агрессивные действия и по-разному на них реагируют. В дальнейшем с использованием процедуры условной вероятности анализировалось влияние на агрессию специфических форм поведения. При этом были выявлены как провоцирующие стимулы, связанные с повышением вероятности нанесения удара, так и подавляющие стимулы, связанные со снижением вероятности нанесения удара. В целом провоцирующие стимулы чаще всего исходили от сиблингов, а подавляющие — от родителей. В категорию подавляющих стимулов входят как позитивные реакции, например, Одобрение или Позитивное физическое воздействие, явно подталкивающие к альтернативному, просоциальному поведению, так и негативные реакции, например, Запрет или Крик. Плач, Смех и Удар — примеры провоцирующих стимулов. Удар оказался наиболее действенным провоцирующим стимулом, что указывает на цикличность и стойкость агрессии в семье.

Стратегия, которой придерживались Паттерсон и Кобб, — одна из форм **последовательного анализа**. Цель последовательного анализа — уловить в потоке социального поведения причинно-следственные связи. Эта процедура предполагает анализ социального взаимодействия, по меньшей мере, двух человек. Наблюдается и регистрируется поведение обоих испытуемых; в одних случаях оно оценивается с помощью одной системы, в других — используются разные системы (как, например, в исследованиях взаимодействия матери и младенца). Какая бы система кодировки ни применялась, фиксируется не только эпизод демонстрирования некоего поведения обоими испытуемыми, но и временная последовательность этих эпизодов. Допустим, нас интересуют мать и младенец, а конкретно связь между материнским прикосновением и улыбкой младенца. Регистрируя последовательность событий, мы можем обнаружить, что частота материнских прикосновений повышается вслед за повышением частоты улыбок младенца — то есть что условная вероятность прикосновения вслед за улыбкой малыша выше исходного уровня. Это означало бы, что направление причинной связи — от младенца к матери, где причиной является улыбка младенца, а следствием — прикосновение матери. Или же мы могли бы обнаружить, что вслед за прикосновением матери повышается частота улыбок младенца, и это означало бы, что прикосновение матери играет роль причины улыбки. Логика последовательного анализа та же, что и логика описанного ранее перекрестного корреляционного анализа с лагом в лонгитюдном исследовании: причина предшествует следствию, и поэтому о причинно-следственных отношениях между двумя переменными судят по тому, насколько изменения одной переменной в период 1 связаны с изменением другой переменной в период 2.

О преимуществах и недостатках натуралистических исследований мы уже не раз говорили, поэтому лишь кратко напомним об основных моментах. Лабораторным процедурам изучения социализации, как правило, свойственна искусственность, а отсюда и невысокий уровень внешней валидности. Рейтинговые методики изучения социализации (тема следующего раздела) представляют собой вербальные отчеты, которые могут отражать, а могут и не отражать фактическое поведение. В натуралистическом наблюдении преодолеваются оба эти недостатка: оно направлено на поведение, а не суждения о нем; кроме того, при наблюдении изу-

чается естественное поведение в естественной среде. Подобные исследования дают уникальную возможность ответить на вопрос, сформулированный Макколл (MacCall, 1977): не только может ли некий социальный опыт оказать влияние, но оказывает ли он его фактически в естественной для ребенка среде? Разработка метода последовательного анализа позволила, используя натуралистический подход, не только получать описание поведения, но и приблизиться к пониманию скрытых причин и следствий.

Обсуждение вышеупомянутого проекта — удачный повод вспомнить о трудностях, которые возникают при проведении натуралистического наблюдения. Главная проблема — получение точных данных. Эта проблема приобретает особое значение, когда предпринимается попытка зафиксировать широкий диапазон поведенческих проявлений, причем в неконтролируемых естественных условиях. Следующая проблема также связана с недостаточным контролем условий — низкая частота эпизодов поведения, интересующего исследователя. Не имея возможности экспериментально вызвать определенные реакции, исследователь может обнаружить, что некоторые формы поведения в период наблюдения проявляются слишком редко, чтобы имело смысл их изучать. Один из важнейших недостатков — влияние наблюдателя на наблюдаемое поведение. Вновь ярким примером служит описанное нами исследование: дело не только в том, что в дом приходят незнакомые люди, но и в том, что цель исследования состоит в фиксировании таких негативных форм поведения, как нанесение ударов или крик — тех форм поведения, которые, как правило, подавляются в присутствии наблюдателя. Наконец, метод последовательного анализа, хотя, несомненно, и информативен, остается по сути корреляционным, а не экспериментальным и, будучи таковым, не может обеспечить точность выводов о причинно-следственных связях, что возможно при использовании экспериментальных манипуляций.

Методики вербального отчета

Мы уже рассмотрели два из основных способов получения информации о процессе социализации: лабораторное исследование и натуралистическое наблюдение. Третий подход — получение информации о методах социализации через вербальные отчеты. Поскольку процесс социализации предполагает наличие, как минимум, двух сторон, существует два пути получения таких отчетов: можно расспросить родителей, как они воспитывают своих детей, или спросить об этом у самих детей. К первому способу, как нетрудно догадаться, прибегают значительно чаще, чем ко второму. Первый пример, который мы рассмотрим, связан с использованием самоотчетов родителей, а второй основан на отчетах подростков о методах социализации, используемых их родителями.

В первом случае исследователей интересовал вопрос о последовательности в своих дисциплинарных воздействиях матерей в ситуациях непослушания ребенка (Grusec & Kuczinski, 1980). Действительно ли — как явствует из множества моделей социализации — матери придерживаются некоего стиля дисциплинирования или реагируют сообразно специфической ситуации? Пытаясь получить ответ на это вопрос, Грусек и Кучински предъявляли своим испытуемым (матерям 4–5 и 7–8-летних детей) 12 типичных ситуаций непослушания ребенка, то есть простые

сценарии, записанные на магнитофон, в которых мать налагала запрет, а ребенок не слушался. Сценарии различались по характеру запрета и неповиновения, а также по последствиям непослушания ребенка. В одних случаях, к примеру, никаких нежелательных последствий не было, в других же за непослушанием следовало нанесение ущерба людям или имуществу. Ниже в качестве примера приведен один из сценариев:

(Слышатся удары мячом)

Мать: Ким, не играй здесь с мячом. Слишком близко к столу. Это дорогая ваза, она принадлежала еще твоей бабушке.

(Слышатся удары мячом)

Мать: Ким, унеси этот мяч куда-нибудь в другое место. Ты собьешь вазу.

(Слышатся удары мячом, затем звук разбивающейся вазы) .

Вопрос, следующий за прослушиванием каждого эпизода, заключался в том, как бы отреагировала мать, если бы это был ее собственный ребенок. Матери могли свободно высказываться по этому поводу, то есть своими словами описывали то, что бы они сделали и сказали, а не выбирали из предлагаемых экспериментатором альтернатив. Затем в этих ответах при помощи системы кодов выделялись проявления разного рода методов дисциплинирования. Среди кодированных категорий было и обоснование предъявляемых требований, которое изучалось в описанных ранее лабораторных исследованиях — к примеру, указание на то, что поведение ребенка огорчит мать. Кодировались также разнообразные формы так называемых методов принуждения — к примеру, физического воздействия на ребенка с тем, чтобы заставить его выполнять требования, или наказание за непослушание. Один из выявленных в исследовании фактов заключался в том, что принуждение оказалось довольно часто используемой формой контроля. Еще один вывод состоит в том, что в рамках одного сценария и в разных сценариях матери используют разнообразные методы, то есть они не всегда последовательны в своих реакциях. Таким образом, специфика ситуации действительно является одной из важнейших детерминант способа реагирования матери.

Второй пример исследования, направленного на изучение методов воспитания, взят из другой исследовательской программы (Dornbusch et al., 1987; Lamborn et al., 1991; Steinberg, Elmen, & Mounts, 1989). Она несколько отличается от подхода Грусек и Кучински (Grusec & Kuczinski, 1980). Информацию о методах воспитания получали не от родителей, а от самих детей — подростков 14–18 лет. Кроме того, в отличие от предыдущего подхода предполагался выбор подростками одного из уровней на каждой шкале ответов. Наконец, исследование проводилось не в форме интервью, а через опросники, заполнявшиеся в процессе группового тестирования в школе.

Дорнбуш и коллеги исходили из теоретических положений о методах воспитания, сформулированных Баумринд (Baumrind, 1971, 1989). Используя сочетание натуралистического наблюдения и родительских отчетов, Баумринд выделила несколько стилей воспитания, которые отличались по таким параметрам, как тепло-

та, контроль и родительское участие. Авторитетный стиль, к примеру, характеризовался строгим контролем в контексте теплых взаимоотношений и частого использования обоснования своих требований, для авторитарного стиля свойственны недостаток тепла и жесткий контроль преимущественно через использование методов принуждения. Один из вопросов, стоявших перед исследовательской группой Дорнбуш, заключался в том, можно ли выделить такие же стили на основе отчетов подростков о способах воспитания, используемых их родителями. Поэтому пункты опросника (которые не всегда совпадают в разных исследованиях) были подобраны так, чтобы с их помощью можно было измерить вариации по таким параметрам, как теплота, контроль и участие. В табл. 12.7 приведено несколько примеров вопросов, использовавшихся в одном из исследований. Главный вывод, к которому пришли исследователи, заключался в том, что типология Баумринд, которая была изначально разработана на основе результатов исследования семей дошкольников, применима и к родителям подростков. Еще один вывод касается наличия корреляционной связи между стилем воспитания и различными показателями компетентности и адаптации подростков. В целом, авторитетный стиль был связан с позитивными результатами (например, высоким уровнем успеваемости), а авторитарный — с негативными результатами. Эти данные подтвердили выводы Баумринд, касающиеся относительной эффективности различных стилей воспитания¹.

Таблица 12.7

Примеры вопросов из «Отчетов подростков о родительском поведении» (Adolescent Report-of-Parenting Scale)

Теплота / участие родителей
Как ты думаешь, что характерно или не характерно для твоего отца (отчима, опекуна)? (Категории ответов: «обычно характерно» и «обычно нехарактерно»)
Я могу рассчитывать на его помощь в любой трудной ситуации.
Он постоянно заставляет меня делать все как можно лучше.
Он помогает мне с учебой, если я что-то не понимаю.
Если он что-то требует от меня, то объясняет, почему.
Часто ли твои родители требуют, чтобы ты усерднее занимался, если ты приносишь домой плохую отметку? (Категории ответов: «никогда», «иногда» и «как правило»)
Часто ли тебя хвалят родители за хорошую отметку? (Категории ответов: «никогда», «иногда» и «как правило»)
Насколько хорошо родители знают твоих друзей? (Категории ответов: «не знают», «плохо знают» и «хорошо знают»)
Строгость / контроль родителей
До какого часа, самое позднее, ты можешь остаться на вечеринке в БУДНИЙ ДЕНЬ (понедельник-четверг)? (Категории ответов: «мне не позволяют оставаться», «до 20:00», «20:00–20:59», «21:00–21:59», «22:00–22:59», «до 23:00 и позже» и «сколько захочу»)

¹ Однако вспомним то, о чем говорилось в главе 6: результаты разных социокультурных групп, включенных в исследование, несколько различались.

До какого часа, самое позднее, ты можешь остаться на вечеринке в ПЯТНИЦУ или СУББОТУ? (Категории ответов: «мне не позволяют оставаться», «до 21:00», «21:00-21:59», «22:00-22:59», «23:00-23:59», «00:00-00:59», «1:00-1:59», «после 2:00» и «сколько захочу»)

Мои родители обычно точно знают, где я нахожусь после школьных занятий. (Категории ответов: «да» и «нет»)

Источник: *Patterns of Competence and Adjustment among Adolescents from Authoritative, Authoritarian, Indulgent, and Neglectful Families*, S. D. Lamborn, N. S. Mounts, L. Steinberg, & S. M. Dornbusch, 1991, *Child development*, 62, p. 1063–1064.

Перейдем теперь к оценке подхода, основанного на вербальных отчетах. Сильные и слабые стороны подхода становятся особенно очевидными при сравнении его с другими подходами. Как мы отметили, лабораторные исследования обладают преимуществом экспериментального контроля, однако нередко цена контроля — искусственность и недостаточные возможности генерализации. Кроме того, лабораторные исследования, даже если в экспериментальной процедуре принимают участие родители, позволяют узнать только о возможном положении вещей, а не о том, что происходит в действительности. Обсервативные исследования позволяют выяснить, что родители делают в действительности, однако с существенной оговоркой: что делают родители, отдающие себе отчет в том, что их наблюдают. Кроме того, результаты обсервативных исследований зависят от способности наблюдателей правильно проинтерпретировать сложные последовательности социальных действий. И лабораторные, и обсервативные исследования охватывают лишь ограниченный диапазон ситуаций и форм поведения.

Именно последнее обстоятельство — широта охвата — составляет, наверное, наиболее очевидное преимущество подхода, основанного на вербальных отчетах. Используя вербальные отчеты, мы приобретаем возможность за сравнительно короткий промежуток времени получить данные, касающиеся широчайшего диапазона методов социализации, значительно более широкого — с точки зрения периодов времени, внешних условий и конкретных форм поведения, — чем диапазон, который можно охватить в натуралистическом наблюдении или в лабораторном исследовании. Эти данные, кроме того, касаются естественного поведения родителей в естественных условиях поведения, не искаженного присутствием наблюдателя или экспериментальными манипуляциями. Хотя непосредственно мы оцениваем вербальный отчет, а не само поведение, отчет дают люди, как никто другой владеющие информацией — иногда уникальной — о родительском поведении: либо сами родители, либо дети, на которых направлено их социализирующее воздействие.

Преимущества подхода, основанного на вербальных отчетах, можно поставить под сомнение одним очевидным, решающим вопросом: точны ли подобные отчеты? В целом, на него можно ответить так: «иногда да, а иногда нет». О том, что вербальные отчеты о социализации не всегда точны, говорят разнообразные данные. Коэффициент корреляции между родительскими отчетами об используемых ими методах социализации и результатами прямого наблюдения родительского пове-

дения, в лучшем случае, умеренный (например, Yarrow, Campbell, & Burton, 1968). В исследованиях, где об одном агенте социализации сообщают сразу несколько человек (например, мать, отец и ребенок — о поведении матери), также корреляционные связи оказываются лишь умеренными (например, Schwarz, Barton-Henry, & Pruzinsky, 1985). Отчеты бывают неточными по разным причинам. В некоторых случаях родители, сознательно или бессознательно, искажают свои ответы с тем, чтобы произвести благоприятное впечатление. Поведение «самолюбивого испытуемого» (см. главу 5) вызывает особое беспокойство, когда речь идет о вербальных отчетах. Родители или дети могут неправильно истолковать вопросы или использовать в своих ответах термины, подразумевая под ними не то, что подразумевает исследователь. Родители могут использовать разные «системы координат». К примеру, две матери могут назвать себя «строгими»; однако для одной быть «строгой» означает время от времени указывать ребенку на его проступки, а для другой это может означать постоянное использование физического наказания. Родители могут просто забыть, как они себя обычно вели или ведут по отношению к ребенку. Проблемы, связанные с забыванием, встают особенно часто в тех случаях, когда используется ретроспективный анализ, то есть изучение методов социализации, применявшихся на более ранних этапах развития ребенка. Результаты исследований свидетельствуют о том, что точность вербальных отчетов о событиях, происходивших несколько лет назад, весьма сомнительна (Yarrow, Campbell, & Burton, 1970).

Методы повышения точности данных, основанных на вербальных отчетах, вытекают из перечисленных выше факторов искажения. Вероятность того, что отчеты будут точными, выше, если задаваемые вопросы будут касаться методов социализации, используемых в настоящем, а не в прошлом. Точность повышает также направленность вопросов на конкретные ситуации и специфические формы поведения (Grusec & Kuczinski, 1980)), минимизирующая затруднения, которые может испытывать респондент, пытаясь разобраться в том, что имел в виду исследователь. Наконец, неизбежное стремление испытуемого проявлять себя с лучшей стороны можно, по крайней мере, ослабить, избегая, где это возможно, оценочных суждений в интервью или опросниках («нет правильных или неправильных ответов», «мы просто хотим понять, что делают родители» и т. д.). Поскольку пробелы в наших знаниях о методах воспитания велики, такие формулировки очень близки к истине.

Точность — одна из важнейших проблем в исследованиях методов воспитания, основанных на вербальных отчетах. Еще одна проблема — выявление причинно-следственных связей. В большинстве исследований методов воспитания ставится цель определить не только, что делают родители, но и как влияют их действия на развитие ребенка. Эту цель ставили перед собой, например, Дорнбуш и др. (Dornbusch et al., 1987, а также Баумринд (Baumrind, 1971, 1989) и ее последователи. Интерес к подобным исследованиям стимулирует возможность нахождения факторов социализации, которые влияют на развитие ребенка либо в положительном (например, авторитетный стиль воспитания), либо в отрицательном (например, авторитарный стиль воспитания) направлении.

Такие исследования неизбежно носят корреляционный характер. Что в них удастся продемонстрировать, так это совместную изменчивость методов воспитания (например, частое использование обоснования своих требований, являющееся одним из компонентов авторитетного стиля) и определенных особенностей ребенка (например, довольно выраженное нравственное поведение). Как отмечалось в главе 3, корреляционный анализ не позволяет выявить основу совместной изменчивости. В исследованиях методов воспитания корреляционные связи всегда имеют несколько альтернативных объяснений. Одно из них: поведение родителей влияет на поведение ребенка — допустим, обоснование родителем своих требований стимулирует нравственное поведение. Это объяснение, несомненно, представляет огромный интерес для исследователя. Второе — поведение ребенка влияет на поведение родителя. С выходом в свет известной работы Белла (Bell, 1968) мысль о том, что дети влияют на своих родителей, приобретает все большее значение в исследованиях социализации. Третий вариант — причинное воздействие в обоих направлениях; возможно, имеет место сложное, реципрокное воздействие поведения родителей на ребенка и поведения ребенка на родителей. Последний вариант — полное отсутствие причинно-следственной связи между определенной формой родительского поведения и тем аспектом развития ребенка, который нас интересует. Возможно, обоснование и нравственное развитие являются результатом действия некоего третьего фактора или совокупности факторов, и связь между обоснованием и нравственным поведением носит лишь статистический, а не причинно-следственный характер.

В главе 3 рассматривались не только разнообразные альтернативные объяснения корреляционных связей, но и методы исключения некоторых вариантов интерпретации. Здесь для нас важно то, что два из описанных методов действительно применяются в первых двух подходах, рассмотренных в разделе «Социализация»: экспериментальные манипуляции с предположительно причинным фактором и определение последовательности событий. Это замечание подводит нас к необходимости еще раз подчеркнуть значение конвергентных операций. Всякий раз, когда отдельные методы изучения некой темы оказываются несовершенными — а «несовершенные», несомненно, подходящее определение для методов, используемых в такой сложной области исследования, как социализация, — единственный путь разрешения проблемы состоит в использовании как можно большего количества различных методов. (Более подробно затронутые в этом разделе вопросы рассматриваются у Bornstein, 1995; Holden & Edwards, 1989.)

Резюме

Эту главу составляют два больших раздела. В первом из них обсуждаются методы изучения важнейших аспектов социального развития ребенка. Во втором — рассматриваются некоторые механизмы социализации ребенка.

Одно из направлений социального развития ребенка — нравственное развитие — включает в себя поведенческий, эмоциональный и когнитивный компоненты. Поведенческий аспект нравственности проявляется в позитивном, просоциальном

действии, к примеру оказании помощи, и в избегании негативных, запрещенных действий, к примеру мошенничества. Нравственное поведение, как и любую проблематику, касающуюся социального развития, можно изучать тремя способами. При использовании натуралистического наблюдения осуществляется попытка оценить естественное поведение в естественных условиях. Лабораторное исследование, напротив, подразумевает создание структурированной ситуации, позволяющей экспериментально актуализировать некую форму поведения. Третий подход — рейтинговая оценка нравственного поведения, которую дают лица, хорошо знающие ребенка, например родители или учителя. Приводятся примеры каждого из этих подходов с последующим обсуждением их плюсов и минусов. Обсуждение сильных и слабых сторон каждого метода заставляет еще и еще раз убедиться в необходимости использования конвергентных операций при изучении вопросов, сложных в методологическом отношении.

В исследованиях эмоционального аспекта нравственности анализируются как негативные эмоциональные проявления, а именно чувство вины, так и позитивные эмоциональные проявления, в частности эмпатия. Чаще всего при изучении чувства вины об эмоциях ребенка судят по тому, какое завершение он придумывает рассказам, в которых описываются разные проступки; при этом предполагается, что ребенок «проецирует» свои чувства на персонажа рассказа. Вербальные методики часто используются и для оценки эмпатии; в этой главе описываются две такие методики, в которых ребенок сообщает о своей эмоциональной реакции на ряд ситуаций. Описываются также исследования, в которых об эмпатии ребенка судят по выражению его лица или физиологическим реакциям на неприятные переживания другого человека.

Инициатором изучения когнитивного аспекта нравственности явился Пиаже. Как и в своих исследованиях когнитивного развития в целом, в работах, посвященных нравственности, Пиаже подчеркивает необходимость использования гибкого клинического метода тестирования и выделяет качественно различающиеся стадии рассуждения о нравственности, которые проходит ребенок в процессе развития. Еще одно, что объединяет работы Пиаже в этих двух направлениях, — это ряд критических замечаний и более глубоких изысканий, для которых они послужили поводом. Работы Пиаже не только стимулировали исследовательскую деятельность, направленную на проверку положений его теории, но и явились предпосылками к формированию основного современного подхода к проблеме нравственного рассуждения, подхода Колберга. Как и Пиаже, Колберг выделяет стадии нравственного развития, основываясь на характере рассуждений детей о нравственных дилеммах; однако и дилеммы, и соответствующие стадии у него значительно сложнее, чем у Пиаже. Сложность системы обработки результатов — одна из трудностей, возникающих при оценке работ Колберга. Спорным представляется также центральное положение теории Колберга о том, что нравственное развитие может быть охарактеризовано некой последовательностью стадий. Обсуждаются разнообразные методы проверки этого положения.

Вторая большая тема, рассматриваемая в этой главе, — развитие Я-концепции. Эта тема, в свою очередь, делится на две подтемы: когнитивный аспект, или вопрос о том, как дети рассуждают о самих себе, и оценочный аспект, то есть вопрос о том,

что думают о себе дети. Изучать когнитивный аспект Я можно уже у детей раннего возраста, при этом анализируется время появления представлений о «персональном влиянии» в период младенчества и визуального самоузнавания в 2–3-летнем возрасте. В исследованиях детей старшего возраста обычно используется метод самоописания, с помощью которого выявляются происходящие в процессе развития изменения характера и уровня сложности представлений детей о самих себе.

Оценочный компонент Я отражается в понятии самооценки. Описывается несколько методик измерения самооценки в детском возрасте, которые основаны на самоотчетах. Некоторые из таких методик являются одномерными, то есть направлены на оценку целостного ощущения собственной ценности; другие — многомерные, то есть направлены на выявление различных аспектов самооценки (например, оценки собственных школьных успехов, успехов в спорте). Результаты исследований свидетельствуют о том, что даже дошкольники обладают некоторой способностью по-разному оценивать себя в различных видах деятельности.

Последний аспект социального развития, который мы рассмотрели в этой главе, — половая типизация и половые различия. Факты, связанные с половыми различиями зачастую оказываются второстепенными по отношению к главным результатам, на получение которых направлены исследования, что существенно затрудняет интерпретацию публикуемых данных о половых различиях и создает риск искажения этой информации. Искажения также возможны на этапе сбора данных; чаще всего, они являются результатом разных ожиданий, которые существуют в отношении мальчиков и девочек. Знание ребенка о поле взрослого также может иметь значение, поскольку дети по-разному реагируют на мужчину-экспериментатора и женщину-экспериментатора.

За обсуждением общих моментов следует рассмотрение конкретных методик, созданных специально для изучения различий между полами и половой типизации. В традиционных тестах половой типизации, например в «Индексе усвоения половой роли», осуществляется попытка измерения соответствия поведения и предпочтений ребенка стереотипам для мужского или женского пола. Недостатком подобных тестов является то, что в них «маскулинность» и «фемининность» рассматриваются как противоположности. Такие методики, как «Полоролевой опросник для детей» включают отдельные шкалы маскулинности и фемининности, что создает возможность «андрогинного» сочетания черт. Наконец, в последнее время все более пристальное внимание уделяется когнитивной стороне половой типизации и представлений ребенка о поле и половых различиях. Здесь интерес представляет способность ребенка определять пол, критерии, используемые им для определения пола, осознание пола как неизменного качества и формирование гендерных схем.

Второй большой раздел этой главы посвящен процессам социализации. Обсуждаются три главных подхода к изучению социализации, которые соответствуют подходам, описанным в первой части главы. В лабораторных исследованиях осуществляется попытка воспроизвести социальный опыт в контролируемой обстановке лаборатории. Рассматриваются примеры, имеющие отношение к двум тщательно изученным вопросам социального развития: влияние модели на поведение ребенка и эффективность обоснования требований желательного поведения. Пре-

имущества лабораторного подхода заключены в экспериментальном контроле, и отсюда возможности выявить причинно-следственные связи. Еще одно преимущество — возможность экспериментальных манипуляций с огромным количеством потенциально значимых переменных. Если говорить о недостатках, лабораторные манипуляции всегда несколько отличаются от своих прототипов реальной жизни, и поэтому всегда необходимо помнить о внешней валидности.

Второй из основных подходов — натуралистическое наблюдение. В этом случае осуществляется попытка измерения социального опыта и его влияния в естественных условиях. Вновь приводится два примера; первый касается реакции детей на полотипизированное поведение их сверстников в детском саду, а второй пример связан с влиянием родителей и сиблингов на агрессию ребенка в домашней обстановке. В натуралистических исследованиях удастся избежать искусственности лабораторных ситуаций и непосредственно зафиксировать поведение в том виде, в котором оно имеет место в естественной обстановке. К проблемам, встающим при проведении подобных исследований, относятся трудности получения точных обсервативных данных, возможность влияния наблюдателя на наблюдаемое поведение и трудности установления причинно-следственных связей в потоке неконтролируемого поведения. Определенные надежды на разрешение последней проблемы связываются с методом последовательного анализа.

Последний из рассмотренных подходов — метод рейтинга, или вербального отчета. В этом случае мы судим о родительском поведении со слов самих родителей или их детей. Описывается два примера этого подхода; в первом исследовании проводится интервью с матерями, во втором используются опросники, которые заполняют дети. Главное достоинство метода вербального отчета заключается в многоплановости получаемой информации. Главный же недостаток — возможность несоответствия вербальных отчетов реальному положению вещей. Рассматривается ряд методов повышения точности отчетов. Обсуждается также еще одна проблема, возникающая при исследовании методов воспитания: интерпретация корреляционных связей между поведением родителей и ребенка с точки зрения причинно-следственных отношений.

Упражнения

В тексте упомянуто исследование (Walker, deVries, & Trevethan, 1987), в котором ответы на гипотетические нравственные дилеммы Колберга сравнивались с ответами на реальные нравственные дилеммы, которые предлагали сами испытуемые. Вопросы, которые задавали авторы этого исследования, предлагая испытуемому рассказать о нравственных дилеммах из реальной жизни, звучали следующим образом:

Попадали ли вы когда-нибудь в ситуацию нравственного выбора, когда вам нужно было принять решение, но вы не были уверены в том, как поступить? Не могли бы вы описать эту ситуацию? В чем для вас в этой ситуации состоял конфликт? Какие обстоятельства вы учитывали, размышляя о том, как поступить? Как вы поступили? Как вы считаете, вы поступили правильно? Почему вы так думаете?

Попробуйте ответить на эти вопросы и предложите ответить на них своим знакомым. Сравните нравственные дилеммы, которые оказались значимыми в вашей жизни, с нравственными дилеммами, которые использовал Колберг.

Проанализируйте вопросы из Шкалы «Отчет о родительском поведении» (Lamborn et al., 1991), представленные в табл. 12.7. Как бы вы ответили на них в подростковом возрасте? Как бы ответили ваши родители, если бы их попросили высказаться по тем же вопросам. Если возможно, задайте эти вопросы вашим родителям и сравните ваши точки зрения. Отражают ли, по вашему мнению, подобные вопросы важнейшие аспекты отношений между родителями и детьми?

В середине 1970-х годов был опубликован отчет об исследовании половой типизации, основанный на анализе перечней предметов, находящихся в комнате ребенка (Rheingold & Cook, *Child Development*, 1975, 46, 459–463). Исследователей интересовал вопрос, существуют ли различия в обстановке комнат мальчиков и девочек — и, как оказалось, они действительно существуют. Ознакомьтесь с этим отчетом и попробуйте воспроизвести их исследование. Посетите как можно больше семей, в которых есть дети, зафиксируйте те же параметры, которые фиксировали Рейнгольд и Кук, и сравните свои результаты с их результатами.

При обсуждении влияния СМИ на развитие детей обычно подчеркивается роль телевидения. Однако книги, которые читает ребенок, также содержат идеи и примеры, влияющие на его представления и поведение. Посетите детский отдел местной библиотеки и наугад возьмите достаточно большое количество книг для маленьких детей. Произведите контент-анализ этих книг в отношении темы, которая вас интересует, отмечая, насколько часто она возникает и как преподносится. Ниже представлены возможные варианты тем.

- Сходства и различия полов.
- Акты насилия и агрессия.
- Изображение различных этнических групп.
- Изображение пожилых людей.

Глава 13

Старение

До недавнего времени термин «психология развития» часто был синонимом «детской психологии». Большинство психологов, именовавших себя специалистами в области психологии развития, в действительности изучали период детства, а большинство работ «о развитии», в действительности, были работами о детях. Но сегодня ситуация несколько иная. Появилось множество работ, охватывающих весь период жизни: это и обычные учебники (например, Kail & Cavanaugh, 1996), и специальные курсы (например, Cohen & Reese, 1994). В связи с расширением границ области психологии развития особый интерес начал вызывать последний отрезок жизни человека, изменения и трудности пожилого возраста. Именно теме пожилого возраста и психологии старения посвящена эта глава.

Пожилой возраст представляет интерес по ряду причин. Самая простая — увеличение числа и доли пожилых лиц в населении. В 1990 году лица в возрасте 65 лет и старше составляли 4 % населения США. Сегодня в эту категорию попадает около 13 % населения, и по оценкам специалистов к 2030 году эта цифра увеличится до 20 %. Кстати, наибольший процентный рост был отмечен в возрастной группе лиц от 80 лет и старше.

Вторая причина интереса к пожилому возрасту сходна с одной из причин интереса к периоду детства. Детство — это во многих отношениях время повышенной уязвимости, когда растущий организм находится «под угрозой» разнообразных нарушений. Изучать детей необходимо в том числе и для того, чтобы узнать, как предотвратить эти нарушения. Преклонный возраст также рассматривается как период уязвимости и риска, как время, когда способности ослабевают и человек сталкивается с множеством потерь (здоровья, работы, супруга или супруги). Как мы увидим, ключевой вопрос при исследовании старения заключается в том, насколько соответствует действительности этот негативный стереотип. Тем не менее, по крайней мере для части пожилых людей проблемы — это очевидная реальность, и необходимость решать их представляет собой еще одно основание для изучения старения.

Третье основание связано с теоретическими соображениями. Считается, «развитие» продолжается всю жизнь, поэтому любая модель развития, конечной точкой которой является подростковый возраст, не может претендовать на полноту. Мож-

но утверждать, что на большом временном отрезке от зачатия до смерти особое внимание следует уделять тем периодам жизни, когда происходят выраженные и существенные изменения. Детство, несомненно, один из таких периодов — время быстрых, неизбежных и глобальных изменений. Многие из этих изменений, кроме того, происходят в одном направлении, отражая движение вперед, прибавление, рост. Пожилой возраст тоже период существенных и, по крайней мере частично, неизбежных изменений, но с существенным отличием: в этом случае, естественное направление изменений — в отрицательную сторону, то есть изменения скорее связаны с потерями, чем с приобретениями. И вновь обоснованность такой негативной характеристики вызывает многочисленные споры. Однако теоретическое значение этой проблемы очевидно, и поэтому возникает необходимость эмпирической проверки.

Наконец, еще одно основание для интереса к старению — следствие многих разногласий и неопределенности, о которых упоминалось выше. Разногласия и неопределенность отчасти обусловлены методологическими трудностями изучения развития в пожилом возрасте и сравнения лиц, находящихся на разных этапах жизненного пути. В действительности, психология старения является наглядным примером того, как трудно провести добротное исследование в целом, и как трудно провести исследование в сфере психологии развития в частности. То есть, она представляет собой интересную и информативную тему из области научной методологии даже для тех, у кого нет непосредственного профессионального или личного интереса к ее изучению.

Здесь для нас, конечно же, важнее всего методологические аспекты исследования старения. Структура этой главы повторяет структуру главы 10, посвященной младенчеству. Глава начинается с рассмотрения общих вопросов изучения старения. Как и в случае с младенчеством, эти вопросы актуальны при изучении любых возрастных групп; однако многие из них имеют особое значение при изучении именно пожилого возраста. Вторая часть главы посвящена специфической проблематике психологии старения.

Общие замечания становятся понятнее на конкретных примерах. Пример, который я использую, — исследования *IQ*. Коэффициент интеллекта — удачный пример не только потому, что это одна из самых популярных зависимых переменных, но и в связи с тем, что многие из методологических вопросов, которые будут обсуждаться, наиболее тщательно анализировались именно в контексте исследования *IQ*. Поэтому в первой части этой главы ставится сразу две задачи: проанализировать методологические вопросы и обсудить результаты исследований важнейшей проблематики.

Общие замечания (на примере исследования *IQ*)

Отбор

Начнем с вопроса, который рассматривался также и в связи с изучением младенцев: как исследователю набрать группу пожилых людей? Ответ тот же, что и в случае с младенцами: разными способами, из которых практически ни один не соответствует модели случайного отбора, предлагаемой учебниками (Camp, West, & Roop, 1989). В отличие от исследователей, изучающих детей и подростков, исследователь, интересующийся старением, не имеет возможности набрать испытуемых

в каком-нибудь специальном учреждении, таком как школа или колледж. Поэтому поиск и случайный отбор из репрезентативной группы пожилых лиц связан с серьезными трудностями. Наверное, наиболее часто используемый способ — обращение в различные группы или организации, в которых состоят пожилые люди, например в религиозные группы, в центры досуга для пожилых людей или в ассоциации пенсионеров. Подобный первичный контингент, очевидно, не обладает достаточной репрезентативностью, поскольку в такие группы, как правило, входят лишь относительно активные и здоровые люди. Выборка может стать еще более искаженной в связи с тем, что участие в исследовании предполагает добровольное согласие; и вновь лишь относительно компетентные пожилые люди соглашаются на участие. Как результат, в исследованиях старения нередко используются в той или иной мере нерепрезентативные выборки пожилых лиц, отличающихся в положительную сторону от популяции в целом.

Получение репрезентативной выборки пожилых людей — это одна из задач исследования старения. Вторая задача — получение выборки пожилых людей, сопоставимой с выборкой молодых людей, с которыми будет осуществляться сравнение. Многие исследователи не ограничиваются простым фиксированием результатов пожилых испытуемых; их цель — сравнить результаты лиц пожилого и молодого возраста. Поскольку в подобных исследованиях в качестве независимой переменной выступает возраст, важно, чтобы сравнивались группы, имеющие как можно больше общего во всех аспектах за исключением возраста. Как указывалось в главе 2, это правило не означает, что исследователь должен пытаться исключить все, чем 20-летние испытуемые отличаются, скажем, от 70-летних. Но оно означает, что необходимо исключить все различия, которые не имеют естественной связи с возрастом.

Отклонение от репрезентативности в одной возрастной группе может исказить результаты ее сравнения с другими группами. Из сказанного выше следует, что лица, принимающие участие в исследовании, обычно отличаются в лучшую сторону от общей популяции. Если молодые испытуемые не будут отбираться, сравнение окажется в пользу пожилых испытуемых. С другой стороны, если группа молодых испытуемых состоит из студентов колледжа (что обычно практикуется), или если пожилые испытуемые отобраны из домов престарелых, искажение будет в противоположном направлении. И даже если во всех возрастных группах используются единые критерии отбора, несоответствие выборок может явиться результатом неодинакового количества лиц, выразивших желание принять участие в исследовании (показатель, который редко включают в исследовательские отчеты о старении).

При сравнении молодых и пожилых лиц особые затруднения вызывают смешения факторов двух типов. Первое связано с различиями в уровне образования. Предположим, проводится сравнение среднего уровня *IQ* у лиц 25- и 75-летнего возраста. Допустим, что исследователю удалось разрешить проблемы, связанные с отбором, о которых говорилось выше, и получить репрезентативные выборки обоих возрастов. Если выборки действительно репрезентативны, они, естественно, будут различаться не только по возрасту, но и по образовательному уровню: в среднем, молодые люди окажутся образованнее пожилых. То есть мы столкнемся со смешением факторов возраста и образования. Заметьте, что складывается парадоксальная ситуация: решение одной из задач отбора, а именно репрезентативности

выборки каждого возраста, в той или иной мере препятствует решению другой задачи, достижению сопоставимости выборок разных возрастов.

В общих чертах проблема, возникающая при исследовании методом поперечных срезов и связанная с различием между представителями разных поколений, рассматривалась в главе 3. Как отмечалось, эта проблема актуальна лишь тогда, когда существует вероятность связи между переменной, значение которой изменяется вместе с возрастом, и изучаемой зависимой переменной. Существует ли вероятность того, что образовательный уровень связан с *IQ*? Разумеется, существует. В действительности, наличие такой связи не только вероятно, но и доказано в ряде исследований (Gonda, 1980). В стандартизированных данных по WAIS, к примеру, коэффициент корреляции между уровнем образования и *IQ* оказался равным 0,68 (Wechsler, 1958). При использовании последовательного анализа (Birgen & Morrison, 1961) было обнаружено, что связь между *IQ* и образовательным уровнем значительно сильнее связи между *IQ* и возрастом.

Что же делать с этим смещением? Как указывалось в главе 3, единственно верного решения нет. Можно при отборе испытуемых привести в соответствие средний уровень образования в разных возрастных группах — к примеру, повысив требования к уровню образования у пожилых людей и несколько понизив их к молодым людям (например, Green, 1969). С помощью этой процедуры обычно удается сгладить нежелательные возрастные различия в *IQ*, но не исключить их полностью. Очевидный недостаток этого приема состоит в том, что при его использовании нарушается репрезентативность выборки. Различия в образовательном уровне можно учесть при статистической обработке данных, в частности, с помощью метода, называемого анализом ковариаций. Однако, как и другие формы статистической подгонки, анализ ковариаций обладает рядом недостатков, а его применение в случае смещения возраст/образование было подвергнуто критике (Storandt & Hudson, 1975). Кроме того, приведение выборок в соответствие по определенному параметру и статистическая подгонка в лучшем случае могут уравнивать испытуемых по *количественным* характеристикам образования, то есть по продолжительности обучения. В этих процедурах не учитывается ни *новизна* образовательных программ, ни *качество* образования.

Различия в уровне образования являются частным случаем эффектов когорты в исследованиях методом поперечных срезов, к которым мы вскоре обратимся при обсуждении планирования исследования. Пока стоит отметить лишь еще один момент. То, насколько серьезные затруднения вызовет взаимосвязь между возрастом и образовательным уровнем, зависит от того, какого рода выводы мы стремимся получить. Если сфера наших интересов ограничивается различиями между молодыми и пожилыми людьми, беспокоиться не о чем; наличие межвозрастных различий в *IQ* — это просто факт, который неизменно подтверждается в многочисленных исследованиях. Проблемы возникают в том случае, когда мы пытаемся перейти от различий к возможности *изменений*. При использовании метода поперечных срезов нельзя получить непосредственное подтверждение снижения *IQ* с возрастом. Интеллектуальному превосходству молодых людей есть альтернативные объяснения, в том числе и различия между поколениями в возможности получить образование. Фактически, рассмотренные выше исследования демонстрируют не только вероятность того, что разница в *IQ* обусловлена различиями в образовательном

уровне, но и то, что эти различия *действительно* объясняют по крайней мере часть межвозрастных различий в *IQ*.

Второе существенное смешение, которое имеет место при использовании метода поперечных срезов для изучения пожилых людей, касается различий в состоянии здоровья. Смешение, как и в предыдущем случае, является неизбежным следствием достижения репрезентативности выборки. В процессе старения здоровье, как правило, ухудшается. Поэтому, получив репрезентативную выборку, мы сталкиваемся с еще одним типом смешения: группы испытуемых будут различаться не только по возрасту, но и по состоянию здоровья.

Связано ли состояние здоровья с *IQ*? Данные ряда исследований говорят о том, что это так. Как и в случае с образованием, при помощи тщательного отбора можно составить группы испытуемых приблизительно соответствующих одна другой по параметру состояния здоровья. Для этого необходимо в выборку пожилых испытуемых включать только лиц с исключительно хорошим здоровьем (например, Botwinick & Birren, 1963). С помощью этой процедуры обычно удается смягчить (но опять-таки не исключить) межвозрастные различия в *IQ*. Отметим, что этот метод страдает теми же недостатками, что и метод уравнивания испытуемых по уровню образования: мы создаем выборку лиц преклонного возраста, в положительную сторону отличающихся от популяции. Можно также проанализировать *IQ* пожилых лиц как функцию от состояния здоровья. Обычно в подобных исследованиях выявляется очевидная связь: пожилые люди со сравнительно хорошим состоянием здоровья интеллектуально превосходят своих ровесников, состояние здоровья которых хуже. На уровне интеллекта может отражаться ряд аспектов физического состояния организма. Среди факторов, которые по крайней мере иногда сказываются на значении *IQ*, называют потерю слуха (Sands & Meredith, 1989), гипертонию (Hertzog, Shaie, & Gribbin, 1978), ишемическую болезнь сердца (Shaie, 1996b), атеросклероз сосудов головного мозга (Spieth, 1965) и депрессию (Cunningham & Haman, 1992).

В некоторых отношениях состояние здоровья — даже более сложная переменная, чем образование. Во-первых, состояние здоровья, в отличие от уровня образования, трудно оценить. Как узнать, насколько здоров испытуемый? В идеальном случае у нас должно быть заключение врача, основанное на результатах полного медицинского обследования. Хотя в некоторых исследованиях и используются медицинские заключения (например, Shaie, 1996b), большинство из них далеки от этого идеала. Чаще всего о состоянии здоровья испытуемого судят с его слов. Методики оценки состояния здоровья, основанные на самоотчете, различаются по своей сложности, от простейших (например, «все испытуемые оценили состояние своего здоровья как нормальное») до значительно более подробных, тщательно разработанных.

Как мы увидели, соответствие самоотчетов действительности всегда вызывает некоторые сомнения. Поэтому важно отметить, что по крайней мере некоторые методики оценки состояния здоровья, основанные на самоотчете, действительно коррелируют с рейтинговыми оценками врачей и поэтому могут рассматриваться как валидные (например, LaRue, Bank, Yarvik, & Hetland, 1979). Однако факт валидности некоторых методик, естественно, не означает, что все методики доверия заслуживают. Более того, во многих исследованиях старения даже не производят

ся попытка оценить состояние здоровья или рассматривать его как одну из переменных (Elias, Elias, & Elias, 1990).

Переменная состояния здоровья сложнее переменной образования, как с концептуальной, так и с методологической точки зрения. Нетрудно представить себе — или даже обнаружить — такие культуры или исторические периоды, в которых уровень образования и возраст не связаны. Эта возможность разграничения возраста и уровня образования говорит о том, что недостаточное образование не является неотъемлемой частью преклонного возраста — то есть что их ковариация является смещением, характерным для нашего времени и для нашей культуры. Возможность независимости состояния здоровья от возраста значительно менее очевидна. Следует ли рассматривать изолированно изменение состояния здоровья, связанное с ним изменение *IQ* и старение само по себе; следует ли рассматривать его как нечто «вторичное», как «отклонение», а не «нормальное старение»? Или же ухудшение здоровья — обязательная составляющая процесса старения?

Как соотносятся между собой состояние здоровья и возраст — спорный вопрос, на который нам вряд ли удастся ответить в этой книге (подробней см. Anstey, Stankov, & Lord, 1993). Замечу, однако, что каково бы ни было решение этой проблемы, для игнорирования снижения *IQ*, связанного с ухудшением здоровья, как некой случайности или артефакта нет никаких оснований. Состояние здоровья действительно *изменяется* с возрастом, а *IQ* действительно *изменяется* вместе с состоянием здоровья. И даже если бы снижение *IQ* было полностью обусловлено ухудшением здоровья (что еще не доказано), оно не исчезло бы и не потеряло своей значимости.

Последнее различие между переменной уровня образования и переменной состояния здоровья касается их значения при разном построении исследования. Обе переменные играют роль в исследованиях методом поперечных срезов, о чем свидетельствуют приведенные выше примеры. При проведении лонгитюдного исследования различия в образовании не имеют большого значения, в том случае если к моменту первого тестирования все испытуемые уже закончили получение образования в учебных заведениях. Различия же в состоянии здоровья, напротив, могут создать определенные трудности, поскольку состояние здоровья, в отличие от уровня образования, с возрастом изменяется. Таким образом, состояние здоровья в отличие от образовательного уровня объясняет некую долю межвозрастных различий и в исследованиях методом поперечных срезов, и в лонгитюдных исследованиях.

Вспомним теперь некоторые общие вопросы отбора, в нашем случае, отбора испытуемых для лонгитюдного исследования. В главе 3 мы отмечали, что выборка в лонгитюдных исследованиях, как правило, несколько специфична, поскольку лонгитюдное исследование предъявляет к испытуемому особые требования. Мы отметили также, что в лонгитюдных исследованиях, как правило, участвуют испытуемые, представляющие одну когорту. Оба эти фактора могут отразиться на выводах относительно стабильности или изменчивости *IQ* с возрастом. Помимо искажений при первичном отборе испытуемых лонгитюдное исследование связано с возможностью избирательного выбывания испытуемых в процессе исследования. В лонгитюдных исследованиях *IQ* это особенно заметно: в целом, наблюдается тенденция выбывания испытуемых с наиболее низкими показателями *IQ*. Такое

**Примеры из раздела «Физическое здоровье» многомерной функциональной оценки
OARS**

№№ во- проса	Вопрос
37 ^a	Сколько раз вы встречались с врачом за последние шесть месяцев, за исключением тех случаев, когда вы лежали в больнице? (Кроме психиатра)
38	За последние шесть месяцев в течение скольких дней вы чувствовали себя так плохо, что не могли заниматься своими обычными делами, например ходить на работу или заниматься хозяйством? 0 Ни одного дня 1 Неделю или меньше 2 Более недели, но меньше месяца 3 1–3 месяца 4 4–6 месяцев – Нет ответа
39	Сколько дней за последние шесть месяцев вы провели в больнице в связи с проблемами со здоровьем?
42	У меня есть список наиболее часто используемых медицинских препаратов. Скажите, пожалуйста, какие из них вы принимали <i>в прошлом месяце</i> ? [Приводится список из 18 лекарств — например, препараты против артрита, обезболивающие, препараты, снижающие артериальное давление]
44	Страдаете ли вы каким-либо из перечисленных заболеваний на сегодняшний момент? [Приводится список из 26 заболеваний — например, артрит, глаукома, астма. Респондент указывает на наличие или отсутствие заболевания, а также на степень его выраженности]
53	Как бы вы оценили состояние вашего здоровья на сегодняшний день — отличное, хорошее, удовлетворительное или неудовлетворительное. 3 Отличное 2 Хорошее 1 Удовлетворительное 0 Неудовлетворительное – Нет ответа
54	Сегодня ваше здоровье лучше, примерно такое же или хуже, чем 5 лет назад? 3 Лучшее 2 Примерно такое же 0 Хуже – Нет ответа
55	Насколько проблемы со здоровьем мешают вам делать то, что вы хотите делать, — абсолютно не мешают, немного мешают или сильно мешают? 3 Абсолютно не мешают 2 Немного мешают 0 Сильно мешают – Нет ответа

^a Раздел «Физическое здоровье» начинается с пункта номер 37.

выбывание создает очевидное искажение в пользу старших возрастных групп: чем старше испытуемые, тем более выражено искажение выборки.

Сиглер и Ботвиник провели исследование, дающее яркий пример феномена избирательного выбывания (Siegler & Botwinick, 1979). Первый раз их испытуемые проходили тестирование в возрасте 65–74 лет; а в течение следующих 10 лет тестирование проводилось до 10 раз, в зависимости от доступности испытуемых. Часть испытуемых выбыла уже после первого тестирования, часть — после первых двух тестирований и т. д., так что до 11 тестирования дошло только 8 из 130 человек. На рис. 13.1 схематически изображены результаты первого тестирования как функция от количества тестирований, на которые пришли испытуемые: заметьте, что разница между лицами, выбывшими после первого тестирования, и теми, кто дошел до конца исследования, составила почти 30 пунктов. Вывод ясен: чем большее количество раз мы тестируем испытуемых в ходе лонгитюдного исследования, тем больше наша выборка будет в положительную сторону отличаться от популяции в целом.

Почему это происходит? Избирательное выбывание имеет два основных объяснения. Первое отражает желание или нежелание людей продолжать участие в исследовании. В целом, испытуемые со сравнительно низкими результатами при первом тестировании менее склонны продолжать участие в исследовании, чем испытуемые со сравнительно высокими показателями. Эту закономерность иллюстрирует исследование Оуэнс (Owens, 1966). Его исходным пунктом было предъявление в 1919 году Армейского альфа-теста (*Army Alpha test*) (одного из первых тестов интеллекта взрослых) 363 первокурсникам университета штата Айова. Была сделана попытка найти и повторно протестировать как можно больше испытуемых из первоначальной выборки, состоявшей из 363 человек. Был найден 201 потенциальный испытуемый; 138 оказались готовы к сотрудничеству — то есть согласились на повторное тестирование, а 63 отказались от сотрудничества. Анализ результатов 1919 года показал, что средний уровень испытуемых, согласившихся на повторное тестирование, составлял 5,63, а средний уровень испытуемых, отказавшихся от участия, составлял 5,20 — статистически значимая разница (она сравнима с разницей около 7 пунктов *IQ* по стандартной шкале *IQ* со средним значением 100). Таким образом, более склонными к повторному тестированию оказались лица с наиболее высокими значениями *IQ*. В других исследованиях (включая упомянутое выше (Siegler & Botwinick, 1979)) констатируется аналогичное избирательное выбывание, причем при интервалах между тестированиями значительно меньших, чем 31 год, как в исследовании Оуэнс.

Вторая причина избирательного выбывания — смерть. В лонгитюдных исследованиях с лицами преклонного возраста некоторые испытуемые умирают в периоды между тестированиями. В среднем у испытуемых со сравнительно низкими тестовыми показателями риск смерти выше, чем у испытуемых со сравнительно высокими показателями (например, Botwinick, West, & Storandt, 1978). То есть низкое значение *IQ* оказалось предиктором — хотя и неточным — приближающей смерти.

Один из особенно интересных примеров связи между *IQ* и продолжительностью жизни — феномен **терминального спада**. Терминальный спад — это довольно резкое снижение умственных способностей за несколько месяцев или лет до смерти. Предположительно он связан с общим ухудшением функционирования организ-

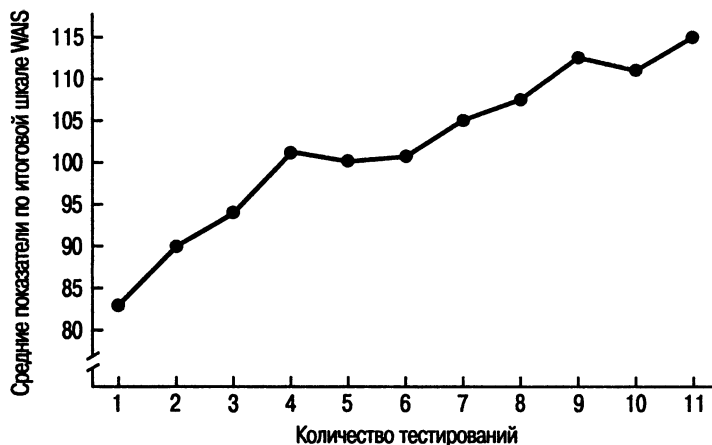


Рис. 13.1. Средние показатели IQ при первоначальной оценке как функция от того, сколько раз испытуемый прошел тестирование. (*Experimental Psychology and Human Aging* (p. 84), D. H. Kausler, 1982, New York: John Wiley & Sons)

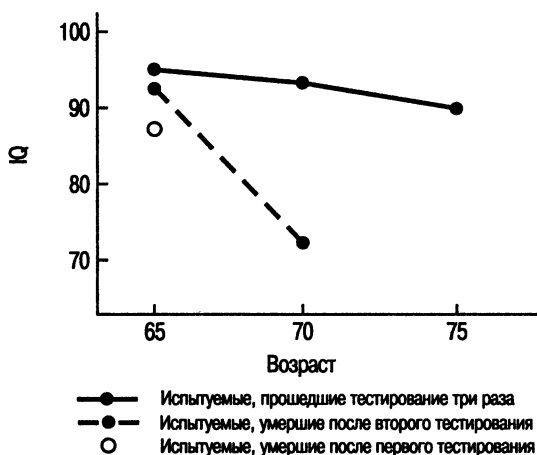


Рис. 13.2. Снижение IQ, отражающее терминальный спад (гипотетические данные)

ма, которое является одним из первых признаков надвигающейся смерти. Терминальный спад можно изучать только в лонгитюдном исследовании, которое включает в себя по меньшей мере три тестирования. Допустим, проводится лонгитюдное исследование с оценкой IQ испытуемых в 65-, 70- и, наконец, в 75-летнем возрасте. Испытуемых в этом исследовании можно разделить на три группы: те, кто пройдет тестирование все три раза, те, кто пройдет тестирование дважды, но не доживет до третьего, и те, кто пройдет только одно тестирование, не дожив до остальных. Нас будет интересовать вторая группа. В среднем уровень интеллекта этих испытуемых падает от момента прохождения тестирования первый раз ко второму. Эта закономерность в графическом виде изображена на рис. 13.2. Терминальный спад состоит именно в ухудшении результатов незадолго до смерти.

(Заметьте, что этот феномен не всегда так ярко выражен, как явствует из рисунка, а то, насколько он типичен и значим, остается спорным вопросом (см. Berg, 1996).)

Возможно, было бы полезно подвести краткие итоги сказанному в этом разделе о связях между состоянием здоровья и *IQ*. Связь между состоянием здоровья и *IQ* действительно существует. В среднем лица с хорошим состоянием здоровья лучше справляются с *IQ*-тестами, чем лица с не очень хорошим состоянием здоровья. Существует также связь между состоянием здоровья и возрастом. В среднем, чем старше люди, тем чаще у них возникают проблемы со здоровьем. Связь здоровья с возрастом имеет некоторые следствия при изучении выборок пожилых лиц. И при использовании метода поперечных срезов, и при использовании лонгитюдного метода выборки лиц самых старших возрастов по состоянию здоровья в положительную сторону отличаются от популяции, что уменьшает межвозрастные различия в *IQ*. Однако даже при этом искажении разница в состоянии здоровья между разными возрастными группами остается, и эта разница, очевидно, объясняет по меньшей мере часть межвозрастных различий в *IQ*.

Измерение

До сих пор мы рассматривали результаты исследования *IQ*, не касаясь важнейшего вопроса: как он измеряется в подобных условиях. Этот раздел посвящен обсуждению методов измерения *IQ*, а также общих вопросов измерения при изучении старения.

Как и для изучения любой другой возрастной группы, для изучения пожилых лиц используется множество разнообразных методик измерения интеллекта. Однако чаще всего применяются и имеют наибольшее значение два теста. Первый — «Векслеровская шкала интеллекта взрослых» (*Wechsler Adult Intelligence Scale*; Wechsler, 1981), или *WAIS*. *WAIS* мы уже кратко рассматривали в главе 11 при обсуждении понятия *IQ* и привели несколько примеров заданий из этого теста (в табл. 11.5). *WAIS*, как и другие тесты Векслера, проводится индивидуально и дает единый показатель общего интеллекта. Поскольку тест делится на вербальную и невербальную шкалы, с его помощью можно получить и более дифференцированную информацию, которую отражают *IQ*-вербальный и *IQ*-невербальный. Вербальная и невербальная шкалы состоят соответственно из шести и пяти субтестов. Таким образом, на максимально специфическом уровне с помощью *WAIS* могут проводиться возрастные сравнения на основе результатов каждого из субтестов.

Вторая методика, наиболее часто используемая в исследованиях *IQ* и старения, — «Тест первичных умственных способностей» (*Primary Mental Ability Test*; Thurstone & Thurstone, 1962), или *PMA*. *PMA* в отличие от *WAIS* может использоваться при групповом тестировании. Как и *WAIS*, он делится на субтесты. В *PMA* входят следующие субтесты: «Вербальное значение», «Пространственный», «Рассуждение», «Числовой» и «Беглость речи». Примеры типичных заданий из некоторых субтестов представлены в таблице 13.2.

Какие проблемы, связанные с измерением, встают при изучении *IQ* и старения? Одна из основных — проблема эквивалентности измерения. Для сравнения *IQ* разных возрастных групп — и с помощью метода поперечных срезов, и с помощью лонгитюдного метода — у нас должны быть равноценные методики измерения *IQ*. Ко-

нечно, легко добиться буквальной или «формальной» эквивалентности: все, что нам нужно, это использовать один и тот же *IQ*-тест для разных возрастных групп. Вопрос в том, действительно ли наш тест измеряет одно и то же у лиц разного возраста; достигаем ли мы не только формальной, но и «функциональной» эквивалентности (Labouvie, 1980). Если тест применим и валиден по отношению к одной возрастной группе больше, чем к другой, результаты межвозрастного сравнения вряд ли будут иметь значение.

Значительная часть психологических методик, используемых при изучении пожилых лиц — не только *IQ*-тесты, — была разработана для изучения более молодых выборок. *IQ*-тесты, как мы увидели, изначально создавались для прогноза успеваемости, и, соответственно, содержание их было ориентировано на академически-вербальные навыки, необходимые для успешного обучения в школе. Для *IQ*-тестов взрослых также всегда была характерна академически-прагматическая направленность. *РМА* создавался на выборках студентов колледжей, а армейский альфа-тест разрабатывался для прогноза успехов юношей на военном поприще. Хотя *WAIS* стандартизировался на выборках лиц в возрасте от 16 до 74, его содержание представляется более тесно связанным с деятельностью в период молодости, а данные по валидации заключаются преимущественно в оценке прогноза успеваемости в академической или профессиональной сферах, имеющие большее значение в годы молодости.

Заметьте, что проблема эквивалентности измерения выходит за пределы необходимости подбора конкретных вопросов для теста. Она намного шире и распространяется на условия и контекст, в которых производится оценка интеллекта. На сегодняшний день, как правило, эти условия представляют собой структурированную лабораторную обстановку и жестко стандартизованные тесты. Можно утверждать, что использование такого подхода действует не в пользу пожилых испытуемых, которые уже давно окончили обучение, не столь привычны к предлагаемым заданиям и не так мотивированы к их выполнению, как молодые. Возможно, компетентность пожилого человека проявилась бы ярче при ее оценке в естественных, привычных для испытуемого условиях, например, при выполнении домашних обязанностей, при совершении покупок в магазине или при объяснении действия игрушки внуку или внучке¹.

¹ Исследователи Мэй, Хэшер и Штольцфус (May, Hasher, & Stolzhus, 1993) предлагают интересный пример, иллюстрирующий, как одна из особенностей лабораторного исследования негативным образом отражается на успешности выполнения испытуемыми различных заданий. В основу этого проекта были положены исследования, свидетельствующие о том, что с возрастом пик когнитивной активности в течение дня сдвигается; при этом молодые люди предпочитают дневные и вечерние часы, а пожилые — утро. В первой фазе исследования с использованием опросника «Жаворонок/Сова» («Morningness-Eveningness Questionnaire») было подтверждено существование этих возрастных различий, что послужило основанием для отбора испытуемых для второй фазы исследования: группы молодых людей с ярко выраженным предпочтением дневного или вечернего времени и группы пожилых людей с ярко выраженным предпочтением утра. Вторая фаза заключалась в выполнении теста на узнавание, выполнявшегося либо утром, либо в предвечернее время. В результатах утреннего тестирования никаких возрастных различий обнаружено не было, тогда как по результатам дневного тестирования молодые испытуемые значительно опережали пожилых. Вывод ясен: назначение тестирования преимущественно на время, неблагоприятное для пожилых людей, приводит к увеличению возрастных различий.

Таблица 13.2

Типы заданий, включенных в Тест первичных умственных способностей (*Primary Mental Ability Test, PMA*)

Вербальное значение					
Инструкция: Найдите слова, которые означают то же, что и первое слово в строке					
<i>ANCIENT</i> (древний)	<i>A. dry</i> (сухой)	<i>B. long</i> (длинный)	<i>C. happy</i> (счастливый)	<i>D. old</i> (старый)	<i>E. sloppy</i> (мокрый)
<i>QUIET</i> (тихий)	<i>A. blue</i> (голубой)	<i>B. still</i> (спокойный)	<i>C. tense</i> (напряженный)	<i>D. watery</i> (водянистый)	<i>E. exact</i> (точный)
<i>SAFE</i> (безопасный)	<i>A. secure</i> (надежный)	<i>B. loyal</i> (преданный)	<i>C. passive</i> (пассивный)	<i>D. young</i> (молодой)	<i>E. deft</i> (ловкий)
Числовой субтест					
Как в цифровом виде написать: одиннадцать тысяч одиннадцать?					
<i>A. 111 B. 1111 C. 11011 D. 110001 E. 111011</i>					
<i>S + S =</i>	<i>A. 1/8</i>	<i>B. 1/4</i>	<i>C. 1/2</i>	<i>D. 1</i>	<i>E. 2</i>
<i>16 × 99 =</i>	<i>A. 154</i>	<i>B. 1584</i>	<i>C. 1614</i>	<i>D. 15084</i>	<i>E. 150084</i>
Рассуждение					
Инструкция: Какая из букв должна стоять за последней буквой в строке?					
<i>cdcd</i>	1. <i>c</i>	2. <i>d</i>	3. <i>e</i>	4. <i>f</i>	5. <i>g</i>
<i>abcabdabeab</i>	1. <i>b</i>	2. <i>c</i>	3. <i>d</i>	4. <i>e</i>	5. <i>f</i>
<i>ambanbaobapba</i>	1. <i>m</i>	2. <i>o</i>	3. <i>p</i>	4. <i>q</i>	5. <i>r</i>

Источник: *SRA Primary Mental Abilities* (p. 1, 6, 11), L. L. Thurstone and T. G. Thurstone, 1962, Chicago: Science Research Associates.

В последние годы приведенные выше доводы заставили ряд исследователей попытаться оценить «повседневный» интеллект пожилых людей, то есть навыки решения повседневных задач (Willis, 1996; Willis & Schaie, 1994). Эти попытки принимали разнообразные формы; здесь я приведу один пример, а затем в разделе «Память» дополню картину еще несколькими примерами исследований. В табл. 13.3 представлены некоторые пункты из методики, известной как «Наблюдаемое выполнение повседневных задач» (*Observed Tasks of Daily Living*; Diehl, Willis, & Schaie, 1995), или *OTDL*. *OTDL* создана для проведения наблюдения в домашней обстановке: проводящий тестирование дает задания и инструкции и фиксирует результаты, а качество решения испытуемыми той или иной задачи затем оценивается по шкалам — от полного выполнения до полной неудачи при выполнении. Тест включает 21 задание из трех сфер деятельности: приготовление пищи, прием лекарств и использование телефона. К этим сферам обращаются во многих методиках оценки «повседневного» интеллекта; кроме того, применяются задания из области ведения финансовых дел, межличностных отношений, покупок и использования транспорта. Методики раз-

личаются не только по содержанию, но и по способам оценки. Решение повседневных задач в домашних условиях (как в *OTDL*) — один из вариантов измерения; возможно также выполнение заданий в лаборатории (например, Marsiske & Willis, 1995), оценка испытуемым собственных способностей справляться с подобными заданиями (например, Fillenbaum, 1985) и оценка этих способностей кем-то, кто знает испытуемого (например, Fillenbaum, 1978).

Обобщить и подвести краткий итог результатам исследований «повседневного» интеллекта довольно трудно; и не только по причине того, что в последние годы таких исследований проводится все больше и больше, но и потому, что результаты этих исследований зачастую противоречивы. В одних исследованиях, к примеру, обнаруживается достаточно тесная связь между показателями подобных методик и стандартными индексами способностей, как, например, *IQ*; результаты других исследований свидетельствуют о независимости этих двух форм компетентности. Аналогично, в одних случаях оказывается, что способности к выполнению повседневных задач с возрастом ухудшаются, тогда как в других случаях приводятся данные о стабильности этих способностей вплоть до глубокой старости (Berg & Klacynski, 1996). Каковы бы ни были конкретные результаты исследований, значение данного подхода очевидно. Подобные исследования обращают наше внимание на навыки, необходимые в повседневной жизни пожилых людей, и позволяют выявить как индивидуальные различия, так и типы компетентности, имеющиеся у пожилых людей, но не учтенные в традиционных методиках измерения *IQ*.

Вернемся теперь к вопросу *IQ*. В исследованиях интеллекта встает еще одна проблема измерения, в некотором смысле даже более фундаментальная, чем проблема эквивалентности измерения. Интеллект, как любой глобальный конструкт, можно определять и измерять разнообразными способами. *IQ*-тесты воплощают в себе одно из операциональных определений интеллекта — далеко не единственное и далеко не полное. Существует ряд других подходов к интеллекту, некоторые из них мы рассмотрим во второй части этой главы. Кроме того, *IQ* сам по себе — не что-то единое и неделимое, а совокупность разнообразных способностей. Выводы о стабильности или изменчивости *IQ*, вполне вероятно, зависят от того, на какой его аспект мы обращаем основное внимание.

Кратко рассмотрим данные, касающиеся последнего из указанных выше вопросов. Это имеет смысл, поскольку конкретный аспект интеллекта, анализируемый в исследовании, оказывается одной из важнейших детерминант результатов межвозрастного сравнения. Наверное, самый достоверный факт в этом отношении связан с сопоставлением вербальной и невербальной шкал *WAIS*. Межвозрастные различия по невербальной шкале выражены значительно ярче, чем по вербальной. Фактически, вербальный интеллект нередко сохранен до глубокой старости. Невербальный интеллект, напротив, имеет тенденцию к раннему и более заметному снижению. Эта закономерность была названа «классическим паттерном старения» (Botwinick, 1984). Его типичный пример представлен на рис. 13.3. Обратите внимание, что это данные сравнения разных возрастов методом поперечных срезов. То же несоответствие показателей вербального и невербального интеллекта обнаруживается и в лонгитюдных исследованиях, однако здесь временные параметры и масштабы ухудшения отличаются от таковых в исследованиях методом поперечных срезов. Мы вернемся к этому вопросу в разделе, посвященном планированию исследования.

Таблица 13.3

Примеры заданий из «Наблюдаемого выполнения повседневных задач» (Observed Tasks of Daily Living)

Приготовление кекса в микроволновой печи	
Экспериментатор вручает испытуемому полуфабрикат кекса с инструкцией по приготовлению, напечатанной на одной из сторон коробки с полуфабрикатом. В этом задании третий пункт инструкции состоит в следующем: «3. Не открывая печи, поддержите пирог в течение 2 минут при высокой температуре; поверните противень на 180° и пекуте еще 2 минуты». Помимо полуфабриката экспериментатор предоставляет испытуемому необходимые кухонные принадлежности, а также модель передней панели микроволновой печи. Эта панель уже использовалась в предыдущих заданиях. Затем экспериментатор дает испытуемому карточку размерами 10,2 × 15,2 см со следующей инструкцией: «Предположим, что тесто уже на противне и печется. Покажите, как бы вы следовали третьему пункту инструкции по приготовлению кекса». Экспериментатор фиксирует действия испытуемого следующим образом:	
Правильные действия	Действия испытуемого
Ставит противень в микроволновую печь.	1.
Нажимает на кнопку «Очистить».	2.
Нажимает на кнопку «Время» и вводит «200».	3.
Нажимает «Старт».	4.
По прошествии 2 минут поворачивает противень на 180°.	5.
Нажимает «Время» и вводит «200».	6.
Нажимает «Старт».	7.
Заполнение коробочки для лекарств	
Испытуемого знакомят с режимом приема лекарств гипотетическим лицом (по имени Пегги Райт) и показывают бутылочки с таблетками для всех шести лекарств, перечисленных в списке. В этом списке следующие три вида лекарств обозначены крестиком: Гигротон, 50 мг. Инструкция: 1 таблетка утром во время еды. Капотен, 50 мг. Инструкция: По 1 таблетке 3 раза в день. Лазикс, 40 мг. Инструкция: По 1 таблетке через день с понедельника по пятницу включительно. Экспериментатор дает испытуемому коробочку для лекарств с четырьмя небольшими отделениями (утро, день, вечер и перед сном) на каждый день недели (с понедельника по воскресенье включительно). Затем экспериментатор дает испытуемому карточку размерами 10,2 × 15,2 см со следующей инструкцией: «Миссис Райт использует эту коробочку для лекарств, которая помогает ей вовремя принимать нужные таблетки. Наполните ее тремя лекарствами, помеченными в списке, в соответствии с инструкциями по их приему».	
Правильные действия	Действия испытуемого
Гигротон на каждый день утром.	1.
Капотен ежедневно утром, днем и вечером.	2.
Лазикс в понедельник, среду и пятницу утром.	3.

Проверка счета за телефонные разговоры	
Экспериментатор дает испытуемому подробный месячный счет (9 страниц) за телефон. Все страницы пронумерованы, а список делится на рубрики.	
Экспериментатор дает испытуемому карточку размерами 10,2 × 15,2 см со следующей инструкцией: «В какие дни, согласно этому счету, были междугородные разговоры с Орегоном? Звонки с телефонной карточки не учитываются».	
Экспериментатор фиксирует действия испытуемого следующим образом:	
Правильные действия	Действия испытуемого
Открывает страницу 6	1.
4 марта	2.
11 марта	3.
23 марта	4.
24 марта	5.

Источник: Everyday Problem Solving in Older Adults: Observational Assessment and Cognitive Correlates, M. Diehl, S. L. Willis, and K. W. Schaie, 1995, *Psychology and Aging*, 10, p. 491.

Разграничение между вербальными и невербальными навыками по *WAIS* имеет много общего с разграничением, которое часто рассматривается в работах, посвященных интеллекту и старению: разграничению между кристаллизованным и текучим интеллектом (Horn, 1982). **Кристаллизованный интеллект** представляет собой знания, постепенно накапливающиеся через опыт — все то, что мы узнаем за годы получения образования, воспитания и аккультурации. Словарный запас, а также все знания общего характера, которые оцениваются в субтесте «Осведомленность» *WAIS* (табл. 11.5), — образец кристаллизованного интеллекта. В отличие от кристаллизованного, **подвижный интеллект** имеет отношение к реакции на новизну и к навыкам решения задач — то есть к базовым навыкам переработки информации, относительно независимым от опыта. Решение задач с аналогиями, складывание «пазлов», задания из шкалы «Рассуждение» *PMA* (табл. 13.2), — все это формы подвижного интеллекта. С возрастом теряются навыки, связанные именно с подвижным интеллектом. Кристаллизованные навыки остаются более или менее сохранными; фактически, в отдельных своих проявлениях кристаллизованный интеллект может улучшаться в течение всего периода взрослости.

Упоминания заслуживает еще один факт, касающийся неодинаковых масштабов ухудшения разных навыков. Задания *IQ*-тестов можно разделить на «скоростные» и «нескоростные». Скоростные задания — это те, в которых важна скорость реакции; на ответ дается ограниченное время, а особенно быстрый ответ может принести дополнительные баллы. Пример скоростного задания — субтест «Составление картинок» из *WAIS* (см табл. 11.5). Нескоростные задания — это те, при выполнении которых скорость ответа не имеет значения; испытуемый может размышлять над заданием столько, сколько ему необходимо, оценивается сам ответ, а не его скорость. Пример нескоростных заданий — вопросы из «Словарного субтеста» *WAIS*. В целом, разница показателей пожилых и молодых испытуемых по ско-

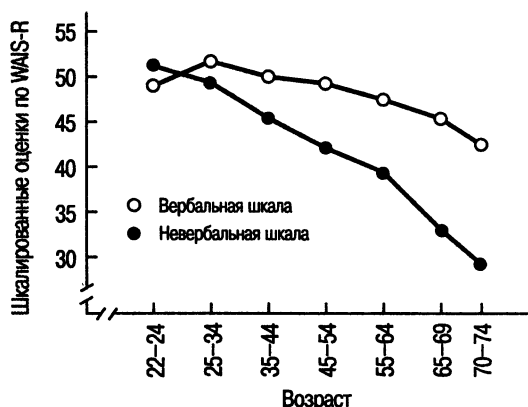


Рис. 13.3. Показатели по вербальной и невербальной шкалам WAIS как функция от возраста. (В конечной стадии показатели невербального интеллекта составляют $\frac{5}{6}$ от показателей вербального интеллекта.) (*Aging and Behavior* (3rd ed., p. 255) J. Botwinick, 1984, New York: Springer)

ростным заданиям больше, чем по нескоростным. Таким образом, для пожилых людей особенно затруднительны ситуации, требующие быстрой реакции.

Большинство данных, касающихся роли скорости при выполнении теста *IQ*, получены с использованием *WAIS*, а это осложняет интерпретацию: все пять субтестов невербальной шкалы *WAIS* скоростные, тогда как вербальная шкала содержит только один скоростной субтест. Таким образом, невозможно определить, что является эффектом содержания задания (вербальный/невербальный), а что — скорости реакции. Важно отметить, что выводы о связи между возрастом и скоростью реакции основываются не только на результатах выполнения субтестов *WAIS*. Снижение скорости реакции с возрастом — факт, подтвержденный исследованиями в разных условиях. Биррен, Вудс и Уильямс, к примеру, заключают, что «наверное, самым типичным и значимым изменением, происходящим в пожилом возрасте, является замедление жизнедеятельности» (Birren, Woods, & Williams, 1980, p. 293). Более того, оказывается, что это замедление обусловлено не только периферическими процессами, к примеру, ухудшением остроты зрения, слуха или потерей ловкости рук. Данные исследования свидетельствуют о том, что, по крайней мере частично, это замедление обусловлено процессами центральной нервной системы и отражает общее замедление скорости переработки информации в мозге, которое сказывается на успешности выполнения разнообразных задач (Salthouse, 1994). Мы вернемся к вопросу о переработке информации ниже в подразделе «Память».

Осталось сделать еще одно замечание общего характера, касающееся измерения. До этого момента мы рассматривали в основном различия между разными возрастными группами в средних тестовых показателях — в частности, различия *IQ* в разные периоды жизни. Хотя среднее — весьма значимый статистический показатель, это не единственная из возможных зависимых переменных. Вполне вероятно, что существуют интересные возрастные различия или сходства, которые сравнение средних не выявляет. Не исключено, что существуют индивидуальные возрастные различия в *изменчивости* показателей — то есть в том, насколько отклоняются показатели от группового среднего. Распространен стереотип, что в процессе старе-

ния индивидуальные различия сглаживаются и люди становятся все более похожими друг на друга. Согласно этому стереотипу, в выборках молодых людей изменчивость должна быть выше, чем в выборках пожилых. В действительности часто бывает все наоборот: если межвозрастные различия все-таки обнаруживаются, у пожилых испытуемых изменчивость оказывается выше. Этот вывод справедлив не только для *IQ*, но и для ряда других зависимых переменных (Nelson & Daneffer, 1992). Поэтому, учитывая только групповые средние, можно упустить из виду интересные различия между группами. Отметим, что указанный выше факт имеет еще одно следствие. Оно касается опасности слишком широких обобщений. Констатация снижения *IQ* (или какого-либо еще показателя) в пожилом возрасте — это констатация снижения *IQ* в среднем по группе. Наличие существенных индивидуальных различий среди пожилых лиц подразумевает наличие значительного количества случаев отклонения от этого среднего.

Изучение изменчивости — это лишь одна из альтернатив изучения средних показателей. Еще один интересный вопрос касается взаимосвязи различных способностей или то, что иногда называют (по образцу такой статистической процедуры, как факторных анализ) «факторной структурой» способностей. Вопрос здесь заключается в том, как соотносятся или коррелируют различные навыки в разном возрасте. Паттерн корреляции может изменяться с возрастом, совершенно независимо от сходства или различия групповых средних. Согласно одной из теорий старения, к примеру, с возрастом граница между различными когнитивными навыками стирается, и они становятся все более взаимозависимыми. Этой «дедифференциации» противопоставляется «дифференциация» навыков, которая, как предполагается, характерна для когнитивного развития в детском, подростковом возрасте и в ранней взрослости. В соответствии с гипотезой дедифференциации корреляционные связи между различными навыками в пожилом возрасте должны усиливаться. Фактически же большая часть данных противоречит этой гипотезе (Schaie, 1996a). Таким образом, вновь общепринятое представление о старении не получает эмпирического подтверждения.

План исследования

Этот раздел начинается с обсуждения двух исследовательских планов, наиболее часто используемых при изучении *IQ* и старения: плана поперечных срезов и лонгитюдного плана. Затем рассматриваются некоторые альтернативные подходы к изучению возрастных изменений.

Выводы, полученные в исследованиях *IQ* методом поперечных срезов и лонгитюдным методом, во многом сходны. Исследования обоими методами свидетельствуют о преимущественно негативном влиянии возраста; пожилые люди справляются с заданиями на интеллект хуже молодых людей. В исследованиях обоими методами на успешность выполнения заданий влияют: состояние здоровья, вербальный/невербальный характер заданий и скорость. Однако есть и некоторые различия. Главное из них заключается в том, что в лонгитюдных исследованиях снижение *IQ* оказывается не столь выраженным, как в исследованиях методом поперечных срезов. Как правило, в лонгитюдных исследованиях снижение *IQ* начинается позже, а его масштабы оказываются меньше, чем при использовании ме-

тода поперечных срезов. Таким образом, лонгитюдное исследование дает более позитивную картину старения.

Чем вызваны эти различия в данных? Два наиболее очевидных объяснения мы уже рассматривали в этой главе и главе 3 при обсуждении общих моментов планирования исследования. В исследовании методом поперечных срезов трудно разделить эффекты возраста и когорты. Молодые и пожилые люди принадлежат не только разным возрастным группам, но и разным поколениям, а различия между поколениями (например, в возможности получить образование) оказываются преимущественно в пользу молодых. Различия между поколениями вполне могут приводить к различиям в *IQ*, независимо от возрастных различий. В лонгитюдном исследовании искажение происходит главным образом в противоположном направлении. В лонгитюдном исследовании имеет место выбывание испытуемых, которое нередко бывает избирательным; при этом в ходе исследования отсеивается все больше испытуемых с низкими показателями *IQ*. Результат — искажение в пользу старших возрастных групп.

Получению результатов, свидетельствующих о снижении *IQ* с возрастом, помимо избирательности выбывания препятствует еще один аспект лонгитюдного исследования. Испытуемые, принимающие участие в лонгитюдном исследовании *IQ*, неоднократно проходят один и тот же тест. Хотя явная обратная связь об успешности выполнения теста и не предоставляется, повторное решение одних и тех же задач вполне может произвести положительный эффект, обусловленный практикой. Данные исследований говорят о том, что эффект практики действительно существует: люди, однажды прошедшие *IQ*-тест, при последующем предъявлении того же теста обычно справляются лучше (Nesselroade & Baltes, 1974). В лонгитюдном исследовании этот фактор может скрыть возрастное снижение показателя *IQ*.

Какие еще методы исследования используются? В главе 3 описывался третий вид «простого» исследовательского плана: план с временным лагом. В исследовании с временным лагом возраст остается константой, а изменяются факторы когорты и времени измерения (см. рис. 3.1). Цель использования такого плана — определение того, влияют ли на изучаемую зависимую переменную факторы, изменяющие свое значение вместе с возрастом, — фактор когорты в исследованиях методом поперечных срезов и фактор времени измерения в лонгитюдных исследованиях.

Описанное ранее исследование Оуэнса (Owens, 1966) — пример сравнения с временным лагом. Как отмечалось, в этом исследовании сравнивались показатели *IQ*, полученные в 1919 году (когда испытуемым было около 19 лет), с показателями *IQ*, полученными при повторном тестировании тех же людей 31 год спустя. Добавим теперь, что в действительности тестирование производилось и третий раз: заключительная проверка в 1961 году, когда испытуемые находились в 61-летнем возрасте. Этот завершающий этап заключался не только в повторном тестировании в рамках лонгитюда, но и в предъявлении того же *IQ*-теста отдельной выборке 19-летних молодых людей. Сравнение показателей 1919 года с новыми показателями выборки 1961 года и составляло часть исследования, проводимую с временным лагом: сравнение двух групп одного возраста, но принадлежащих к разным поколениям и тестируемых в разное время. В исследовании Оуэнса было обнаружено, что показатели новой выборки 19-летних испытуемых были значительно

выше показателей первой выборки, полученных в то время, когда испытуемые находились в 19-летнем возрасте. Этот результат свидетельствует о том, что на уровень выполнения *IQ*-тестов может влиять как время измерения, так и фактор когорты. Вполне вероятно, что имеют значение оба фактора.

Главная альтернатива поперечных и лонгитюдных планов — использование одного или нескольких *последовательных* планов, описанных в главе 3. Как отмечалось, последовательные планы представляют собой сочетание поперечных срезов, временного лага и лонгитюдного подхода. Такие исследования позволяют получить больше информации, чем исследование с применением простых методов, но они же требуют больших затрат времени и усилий. Поэтому сегодня они используются нечасто. Здесь мы рассмотрим широкомасштабное исследование, результаты которого оказались весьма значимыми — исследование Шайе началось в 1956 году с тестирования 500 испытуемых в возрасте от 21 до 70 лет. С целью анализа данных выборка (которую я буду называть выборка 1) была разбита на 7 возрастных групп с разницей в возрасте 7 лет. Средний возраст группы самых молодых на момент первого тестирования составлял 25 лет, а средний возраст старшей группы — 67 лет. Хотя использовалось несколько методик, главной зависимой переменной были показатели *PMA*: как результаты отдельных субтестов, так и итоговые *IQ*.

Тестирование 1956 года представляло собой исследование методом поперечных срезов, результаты которого можно найти в отчете Шайе (Schaie, 1958). Как и в других исследованиях методом поперечных срезов, в этом исследовании вырисовывалась негативная картина интеллекта в старости: из полученных данных следовало, что снижение *IQ* начинается примерно с момента достижения 50-летнего возраста. Однако поперечный срез явился лишь начальным этапом исследования Шайе. В 1963 году было проведено второе тестирование с привлечением двух разных выборок. Одна из них уже проходила тестирование; из прежних 500 она включала 303 испытуемых, которые в 1963 году были протестированы повторно. Теперь, естественно, возрастной диапазон этой выборки составлял от 28 до 77 лет. Вторая выборка (выборка 2) была новой; это была самостоятельная выборка с возрастным диапазоном от 21 года до 74 лет, изучавшаяся методом поперечных срезов. В нее входило 977 испытуемых. Получив данные 1963 года, Шайе с коллегами произвели ряд аналитических процедур и сравнений. Сравнение происходило в трех направлениях и охватывало практически весь период взрослости: выборка 1 в 1956 году, она же 7 лет спустя в 1963 году и выборка 2 в 1963 году. Проводилось также лонгитюдное сравнение с временным интервалом 7 лет: в выборке 1 сопоставлялись результаты 1956 и 1963 годов. Кроме того, производилось сравнение с временным лагом, при котором сопоставлялась каждая из возрастных групп 1956 года с соответствующей возрастной группой 1963 года.

Тестирование 1963 года, несмотря на всю его масштабность и трудность, оказалось отнюдь не последним этапом исследования. В действительности это была лишь вторая «волна» в проекте, который на сегодня включает в себя 6 отдельных «волн», или периодов сбора данных. Шайе с коллегами продолжают сбор данных с 7-летними интервалами, последнее тестирование происходило в 1991 году. Общая стратегия на каждой фазе тестирования осталась неизменной с 1963 года — года первого ретеста. Во-первых, в каждую волну сбора данных в рамках лонгитюдного

Таблица 13.4

Схема последовательного исследования Шайе: средний возраст и объем выборки на каждом этапе тестирования

	Время тестирования					
	1956	1963	1970	1977	1984	1991
Вы- борка 1	25–67 (<i>N</i> = 500) первый тест	32–74 (<i>N</i> = 303) второй тест	39–81 (<i>N</i> = 162) третий тест	46–88 (<i>N</i> = 130) четвертый тест	53–90 (<i>N</i> = 97) пятый тест	60–95 (<i>N</i> = 71) шестой тест
Вы- борка 2		25–74 (<i>N</i> = 997) первый тест	32–81 (<i>N</i> = 420) второй тест	39–88 (<i>N</i> = 337) третий тест	46–95 (<i>N</i> = 225) четвертый тест	53–95 (<i>N</i> = 161) пятый тест
Вы- борка 3			25–81 (<i>N</i> = 705) первый тест	32–88 (<i>N</i> = 340) второй тест	39–95 (<i>N</i> = 223) третий тест	46–95 (<i>N</i> = 175) четвертый тест
Вы- борка 4				25–81 (<i>N</i> = 609) первый тест	32–88 (<i>N</i> = 295) второй тест	39–95 (<i>N</i> = 201) третий тест
Вы- борка 5					25–81 (<i>N</i> = 629) первый тест	32–88 (<i>N</i> = 428) второй тест
Вы- борка 6						25–81 (<i>N</i> = 693) первый тест

анализа осуществляется попытка повторного тестирования как можно большего количества участников предыдущих обследований. Как и следовало ожидать, с увеличением числа интервалов между обследованиями и возраста участников для повторного тестирования удастся привлечь все меньше испытуемых; к 1991 году удалось привлечь лишь 71 испытуемого из исходной группы, состоявшей из 500 человек. Тем не менее лонгитюдные данные за 35-летний период даже части выборки — большая редкость и являются ценным источником информации. Во-вторых, в каждую волну тестирования включалась дополнительная выборка, позволяющая провести поперечный срез с группами испытуемых в возрасте от 20 до 80 лет. В 1991 году, к примеру, было привлечено 693 новых испытуемых, что увеличило число участников проекта до более чем 5000 человек.

Схема проведения этого весьма сложного исследования могла бы помочь разобраться в его подробностях. Такая схема представлена в табл. 13.4. В ней можно выделить ряд возможностей для поперечных, лонгитюдных сравнений и сравнений с временным лагом. Сопоставление этой таблицы с рис. 3.2 и 3.3 поможет увидеть отдельные формы последовательного анализа, включенные в общий план

исследования. Любая пара строк таблицы, к примеру, представляет собой основу для плана последовательности когорт — то есть двух накладывающихся друг на друга лонгитюдных исследований, в которых в качестве независимых переменных рассматриваются возраст и когорта.

Каковы были результаты этого удивительного по своим масштабам исследования? Значительную часть результатов, полученных в исследовании Шайе, мы уже рассматривали, поскольку они легли в основу выводов об *IQ* и старении, приведенных выше. Один из результатов, не вызывающих сомнений, к примеру, касается различия между паттернами изменения разных навыков, оцениваемых *PMA*. По словам Шайе (Schaie, 1994, p. 306), «не существует какого-то единого паттерна возрастных изменений интеллектуальных навыков»; некоторые навыки остаются сохраненными или с возрастом даже улучшаются, тогда как другие в разной степени и начиная с разных моментов жизненного пути ухудшаются. Еще один результат был получен при сопоставлении данных, полученных методом поперечных срезов и лонгитюдным методом. Как явствует из табл. 13.4, сравнение любых возрастных групп (скажем, 60-летних и 74-летних) производится множество раз, как в рамках поперечного среза, так и в лонгитюде; при этом в первом случае неизменно обнаруживаются большие возрастные различия, чем в последнем. Третий вывод связан с предыдущим: когорта играет существенную роль при оценке многих навыков — несомненно, большую, чем считалось до исследования Шайе. Хотя результаты сравнения когорт обычно свидетельствуют в пользу младших поколений, интересно отметить, что были и исключения. Результаты по ряду субтестов, к примеру, достигали максимума у когорты 1924 года рождения, а затем неуклонно снижались.

Нет необходимости подчеркивать, что мы привели лишь часть результатов, полученных в этом широкомасштабном исследовании (в действительности, ряд зависимых переменных и соответствующих теоретических вопросов мы даже не упоминали). Не в полном объеме рассмотрены и сложности анализа последовательных исследовательских планов, которые влекут за собой ряд противоречий и сомнений в отношении того, что же именно показало исследование Шайе. Обсуждение этих вопросов можно найти в приведенной выше литературе (см. также Donaldson & Horn, 1992; Salthouse, 1991; Hertzog, 1996; Hertzog & Dixon, 1996).

Имея представления об общих вопросах, перейдем к специфическим темам исследования старения. *IQ*, конечно же, одна из таких тем. Ниже мы рассмотрим еще один подход к изучению когнитивной сферы в пожилом возрасте, после чего перейдем к обсуждению методов изучения личности и социального развития.

Память

Как отмечалось ранее, *IQ* является одной из популярнейших зависимых переменных психологических исследований старения. Память — еще одна популярная переменная. В 1991–1993 годах 34 % статей, опубликованных в журналах *Psychology and Aging* и *Journal of Gerontology: Psychological Science*, так или иначе было связано с темой памяти (Smith, 1996).

Как и в отношении интеллекта, в отношении памяти в пожилом возрасте существует стереотип. Опять-таки стереотип ухудшения: люди преклонного возраста

уже не способны помнить так же хорошо, как в молодости. Как и в случае с *IQ*, этот стереотип имеет под собой основание, однако есть и некоторые специфические моменты и исключения. Цель большинства исследований памяти в пожилом возрасте состоит в выявлении условий, при которых возникают или не возникают проблемы, связанные с памятью.

Узнавание и воспроизведение

Основное разграничение, проводимое при исследовании памяти в любом возрасте, — разграничение между *узнаванием* и *воспроизведением*. Как отмечалось в главе 11, способность к узнаванию присутствует и довольно развита уже в раннем возрасте, а в способности к воспроизведению наблюдаются более выраженные изменения, связанные с развитием в период детства. Как показали Шонфилд и Робертсон (Schonfield & Robertson, 1966), в способности к воспроизведению также наблюдаются более серьезные изменения в процессе старения. Исследователи предъявляли для запоминания двум возрастным группам испытуемых, в возрасте от 20 до 29 лет и в возрасте от 60 до 75 лет, два списка по 24 слова каждый. Память проверялась двумя способами: испытуемых просили перечислить все слова, которые они смогли запомнить из первого списка (тест на воспроизведение), а каждое слово из второго списка нужно было выбрать из пяти альтернативных вариантов (тест на узнавание). Неудивительно, что результаты выполнения теста на узнавание оказались выше результатов выполнения теста на воспроизведение, причем в обеих группах. Однако разница в уровне выполнения теста на узнавание и теста на воспроизведение в группе пожилых испытуемых была значительно больше. В действительности по результатам теста на узнавание между группами не было выявлено практически никаких различий. Только по уровню воспроизведения молодые испытуемые (среднее — 13,78 воспроизведенных слов) превосходили пожилых (среднее — 7,50 слов). Обратите внимание, что этот факт отражает взаимодействие возраста и экспериментальных условий: эффект возраста изменялся как функция от типа изучаемых мнемических навыков.

Описанное исследование часто приводят в качестве примера, вероятно, потому, что это одна из первых попыток изучить вопрос узнавание/воспроизведение, и потому, что его результаты оказались вполне однозначными. Со времени проведения этого исследования неоднократно проводилось сравнение узнавания и воспроизведения в разном возрасте. Аналогичное взаимодействие возраста и экспериментальных условий выявляются достаточно часто, чтобы можно было с уверенностью говорить о том, что для пожилого возраста затруднения с воспроизведением более типичны, чем затруднения с узнаванием. С другой стороны, далеко не всегда оказывается, что с возрастом способность к узнаванию остается неизменной, как сообщают Шонфилд и Робертсон. В ряде исследований были получены результаты, свидетельствующие о преимуществе молодых испытуемых как по тестам на воспроизведение, так и по тестам на узнавание (например, McCormack, 1981).

Почему способность к узнаванию и способность к воспроизведению неодинаково изменяются с возрастом? Сначала необходимо выяснить, почему воспроизводить вообще труднее, чем узнавать. Одно из очевидных различий между двумя этими мнемическими процессами имеет отношение к различию между хранением

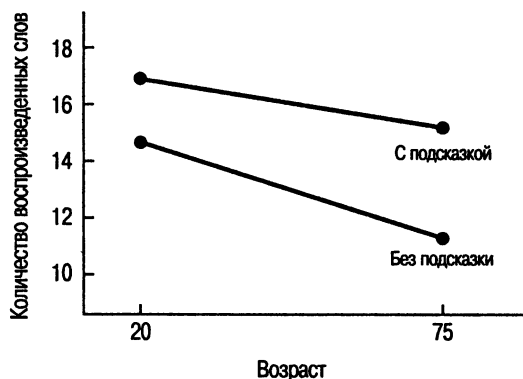


Рис. 13.4. Взаимодействие возраста и экспериментальных условий в исследовании воспроизведения по подсказке («Memory Loss with Age: A Test of Two Strategies for Its Retardation», M. W. Laurence, 1967, *Psychonomic Science*, 9, pp. 209–210)

в памяти и извлечением из памяти. Оба мнемических процесса подразумевают хранение информации, однако для воспроизведения требуется больше усилий для извлечения информации, чем для узнавания, когда правильный ответ уже присутствует. Малая подверженность способности к узнаванию возрастным изменениям подразумевает то, что такой компонент памяти, как хранение практически не затрагивается процессом старения; возрастные изменения по большей части являются результатом ухудшения способности к извлечению информации из памяти. Как и все обобщающие выводы о возрастных различиях в памяти, это обобщение имеет определенные ограничения и исключения (как, например, возможность наличия возрастных различий и в способности к узнаванию). Тем не менее оно действительно имеет под собой основу и объясняет часть возрастных различий в памяти.

Как еще можно изучать разницу между хранением информации в памяти и ее извлечением? Довольно часто используется метод *воспроизведения по подсказке*. В исследованиях воспроизведения по подсказке сравниваются результаты выполнения стандартных тестов на воспроизведение и тестов на воспроизведения, в которых делается намек, дается «подсказка». К примеру, в одном из исследований (Laurence, 1967), тест на воспроизведение проходили молодые (средний возраст 20 лет) и пожилые (средний возраст 75 лет) испытуемые. Стимулами служили 36 слов, соответствующих 6 категориям: цветы, деревья, птицы, неживая природа, овощи и страны. Одной половине испытуемых перед тестированием были выданы карточки с подсказками — названием 6 категорий и указанием на то, что каждое слово принадлежит к одной из них; другая половина испытуемых никаких подсказок не получала. Результаты этого эксперимента представлены на рис. 13.4. Как можно заметить, имело место взаимодействие возраста и экспериментальных условий: пожилым испытуемым подсказка помогала больше, чем молодым. В действительности в ситуации подсказки группы практически не отличались по количеству воспроизведенных слов.

Как и результаты исследования, рассмотренного ранее (Schonfield & Robertson, 1966), при дальнейшем изучении этого вопроса результаты работы, о которой пишет Лоренс, не всегда находили подтверждение. Взаимодействие возраста и экс-

периментальных условий обнаруживалось далеко не во всех исследованиях воспроизведения с подсказкой; в ряде исследований молодым испытуемым подсказка помогала ничуть не меньше, чем пожилым. Общий вывод, который можно сделать, тот же, что и в случае со сравнением узнавания и воспроизведения. То, что подсказка при извлечении информации из памяти особенно помогает пожилым людям, наводит на мысль о том, что пожилые люди неплохо запоминают и хранят информацию, однако испытывают трудности при восстановлении ее в памяти. Тот факт, что подсказка не всегда устраняет различие между пожилыми и молодыми испытуемыми, говорит о наличии других факторов (помимо трудностей при извлечении информации из памяти), определяющих возрастные различия в памяти.

Мнемонические стратегии и конструктивная память

Довод о необходимости активности при воспроизведении материала является лишь частью гипотезы о том, что для пожилых лиц особые трудности представляют ситуации, требующие активной переработки информации — ситуации, в которых, чтобы запомнить, сохранить в течение какого-то времени и при необходимости восстановить информацию, необходимо что-то *сделать* (Craik, 1994). Эта гипотеза должна быть вам знакома, поскольку она была предложена для объяснения некоторых изменений в период детства, — это идея о том, что старшие дети запоминают и помнят лучше младших не только потому, что у них больше объем памяти, но и потому, что они более адаптивно перерабатывают информацию. В отношении детей эта гипотеза проходила проверку в двух основных направлениях, в которых проводятся исследования возрастных различий и в период взрослости: изучение мнемонических стратегий и изучение конструктивной памяти.

Исследования мнемонических стратегий в период взрослости проводятся по тем же методологическим направлениям, что и исследования стратегий в детском возрасте, рассмотренные в главе 11. В некоторых исследованиях об использовании или неиспользовании стратегии судят по успешности выполнения тестов памяти. Об использовании стратегии группирования, к примеру (табл. 11.7), судят по тому, наблюдается ли при воспроизведении тенденция объединять концептуально близкие слова. Аналогично, об использовании стратегии повторения судят по наличию эффекта первичности — то есть лучшего запоминания первых элементов ряда. Еще один подход — экспериментальное побуждение к использованию стратегий. Например, испытуемых можно попросить повторять заучиваемый материал, а затем проверить, как повторение отразилось на успешности выполнения теста.

Многие выводы о роли мнемонических стратегий в пожилом возрасте мы уже рассматривали. Вспомним, что одним из видов прикладных исследований, о которых шла речь в главе 6, является оценка эффективности программ обучения мнемоническим стратегиям. Основанием для использования таких программ служит предположение о том, что некоторые (хотя, конечно же, не все) мнемонические трудности, с которыми сталкиваются пожилые люди, — следствие того, что они не используют эффективные мнемонические приемы. Фундаментальные исследования, описанные выше, подтверждают это предположение. Хотя результаты исследований не всегда однозначны (см. обзор литературы по этой теме (Kausler, 1994)), ряд данных свидетельствует о том, что пожилые люди реже спонтанно используют

мнемонические стратегии, чем молодые, а также реже выбирают оптимальную стратегию. Исследования, в которых испытуемых побуждают к использованию стратегий — как в рамках непродолжительного эксперимента, так и в длительных обучающих исследованиях, описанных в главе 6, — также подтверждают эту гипотезу. Как мы увидели, обучение пожилых лиц использованию стратегий может улучшить их показатели по тестам памяти. В некоторых исследованиях, (например, Canestrari, 1968) было обнаружено, что обучение использованию стратегий помогает пожилым испытуемым больше, чем молодым, то есть имеет место взаимодействие возраста и экспериментальных условий. Этот факт согласуется с гипотезой о том, что молодые люди более склонны к самостоятельному изобретению стратегий, чем пожилые люди, а следовательно, и не так нуждаются в помощи экспериментатора.

Помимо мнемонических стратегий в главе 11 рассматривалась еще одна важная область исследования — конструктивная память. Как отмечалось, конструктивная память имеет отношение к влиянию на мнемические процессы системы знаний — тенденции корректировать информацию, делать выводы, выходя за пределы непосредственно данного, для того чтобы *понять* воспринимаемое. В исследованиях мнемонических стратегий в основном изучается материал, лишенный смысла, и конкретные приемы запоминания, сохранения и воспроизведения информации. В исследованиях же конструктивной памяти используется материал, имеющий смысл, и изучается влияние неспецифической, зачастую неосознанной когнитивной деятельности на мнемические процессы.

Конструктивная память давно является одной из излюбленных тем в литературе, посвященной памяти в период взрослости. Основные результаты исследований касаются сходства мнемических процессов в различных возрастных группах. У взрослых лиц всех возрастов, включая и пожилых, наблюдаются разнообразные проявления конструктивной памяти, которая, по-видимому, качественно не зависит от возраста (Salthouse, 1982). Тем не менее в некоторых исследованиях обнаруживаются возрастные различия. Причем эти различия носят тот же характер, что и другие различия, описанные в этом разделе: превосходство молодых над пожилыми. Рассмотрим один пример; более подробно эту тему раскрывает Кауслер (Kausler, 1994).

Исследование, которое я приведу в качестве примера, связано с памятью на рассказы и поэтому имеет много общего с некоторыми из исследований, описанных в главе 11. В серии экспериментов (Cohen, 1979, 1981) сравнивалась память молодых и пожилых людей на отрывки текста. Примеры двух отрывков приведены во вставке 13.1. Независимо от характера стимулов, интерес представляют два вида памяти. Во-первых, память на информацию, непосредственно отраженную в тексте. Первый вопрос, следующий за каждым отрывком, был направлен на оценку именно дословного запоминания. Второй вид памяти — память на информацию, не представленную непосредственно в тексте, к которой испытуемый должен прийти логическим путем. Второй вопрос, следующий за каждой историей, направлен на оценку запоминания информации, логически следующей из текста. Главный результат, полученный в этом исследовании, заключался в том, что возрастные различия в памяти на подразумеваемую информацию более выражены, чем возрастные различия в памяти на непосредственно данную информацию. Пожилые испы-

Вставка 13.1

Примеры отрывков и вопросов в исследовании конструктивной памяти

В хорошую погоду миссис Браун обязательно ходит на прогулку в парк. Ей нравится наблюдать за игрой детей и кормить уток хлебными корочками. Ей нравится бродить по парку и прогулка от парка до дома. Последние три дня все время лил дождь, хотя на дворе середина лета и город полон отдыхающими.

Чем кормит уток миссис Браун?

Ходила ли вчера в парк миссис Браун?

На первом этаже находятся три комнаты; кухня, столовая и гостиная. Гостиная расположена на стороне фасада, а кухня и столовая выходят окнами на огород на заднем дворе. В комнатах, выходящих на улицу, постоянно слышится неприятный шум от проезжающего транспорта. Мать готовит на кухне, а бабушка читает газету в гостиной. Дети в школе и будут дома не раньше пяти часов.

Что делает мать?

Кому мешает шум от проезжающего транспорта?

Источник: Language Comprehension in Old Age, G. Cohen, 1979, *Cognitive Psychology*, 11, p. 416.

туемые хорошо понимают и запоминают непосредственное содержание отрывка, однако сталкиваются с трудностями при переходе от непосредственно данного к подразумеваемому. А этот переход от непосредственной данности к логическим выводам и составляет суть конструктивной памяти.

Почему существуют такие возрастные различия в конструктивной памяти? У детей часть этих различий объясняется разницей в запасе знаний: старшие дети знают больше младших, поэтому их логика будет отличаться от логики младших детей. Для взрослых же такое объяснение представляется менее правдоподобным; взрослый человек любого возраста, несомненно, в состоянии установить логические связи, чтобы дать ответ в заданиях, приведенных выше. Различие, видимо, обусловлено тем, насколько испытуемые склонны устанавливать такие связи — использовать свои знания с тем, чтобы прийти к определенным выводам. С возрастом люди все менее склонны искать логические связи. Одно из объяснений ухудшения способности к самостоятельному поиску логических закономерностей связано со снижением скорости обработки информации. Утверждается (например, Salthouse, 1982), что у пожилых людей процесс усвоения новой информации протекает медленнее и что более низкие темпы переработки информации мешают направлять внимание на новый материал и устанавливать связи между новым и старым, хранящимся в памяти материалом. Поскольку логические построения основываются именно на связях между новым и старым, происходит ослабление конструктивной памяти. Это объяснение вполне согласуется с результатами других исследований, в которых было продемонстрировано, что снижение скорости реакции в пожилом возрасте негативно отражается на разнообразных аспектах психической деятельности.

«Повседневная» память и долговременная память

Рассмотрим еще два вопроса, имеющих отношение к теме «Память». До сих пор основное внимание уделялось стандартным лабораторным методам, часто исполь-

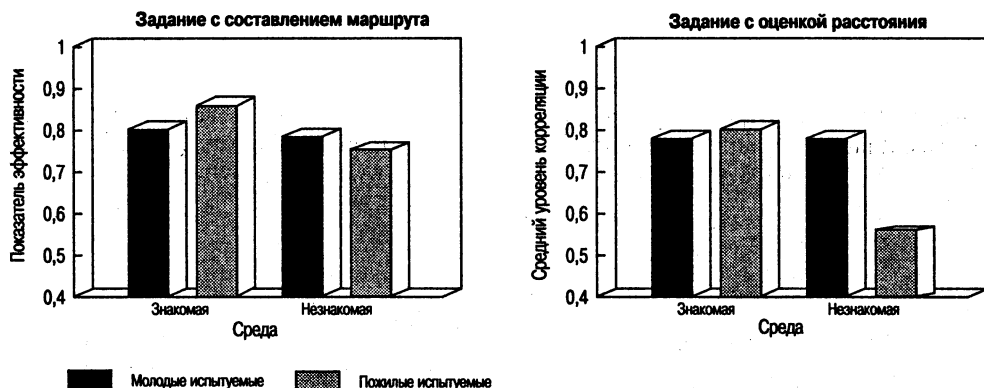


Рис. 13.5. Взаимодействие возраста и контекста в исследовании пространственной памяти Кирасик. Зависимой переменной в задании на составление маршрута было соотношение кратчайшего и выбранного пути; а в задании с оценкой расстояния — корреляция между оцененным и реальным расстоянием. В обоих случаях более высокие значения указывают на более успешное выполнение задания. («Spatial Cognition and Behavior in Young and Elderly Adults: Implications for Learning New Environments», K. C. Kirasic, 1991, *Psychology and Aging*, 6, 10–18, p. 14)

зуемым при изучении памяти в любой возрастной период. Два последних подхода, которые мы рассмотрим, несколько более специфичны; кроме того, при их (независимом) использовании удастся получить более позитивную характеристику памяти у пожилых людей.

Ранее мы отметили, что «повседневный» интеллект стал весьма популярной темой исследования. То же можно сказать и о «повседневной» памяти. Как и в случае с интеллектом, в исследованиях повседневной памяти существует множество направлений и методов. Я приведу лишь два примера (подробней см.: Hess & Pullen, 1996 и West & Sinnott, 1992).

Изучение пространственной памяти — одно из популярных направлений исследования повседневной памяти. Все мы ежедневно и многократно используем те или иные формы пространственной памяти — следуя обычным маршрутом, вспоминая местонахождение различных предметов и т. д. Ухудшается ли такая память с возрастом? Мы уже рассмотрели один из примеров лабораторного подхода к этому вопросу — исследование Черри и Парк (Cherry & Park, 1993), описанное в главе 2. На основе результатов этого исследования было сделано два вывода, достаточно типичных для подобных исследований: пространственная память пожилых испытуемых оказалась хуже пространственной памяти молодых испытуемых, и пожилым испытуемым запомнить местонахождение предметов помогал хорошо различимый, привычный контекст.

В исследованиях психических функций в повседневной обстановке понятие привычного контекста значительно расширяется. Лабораторные условия и экспериментальные аналоги (например, объемная модель и схема в исследовании Черри и Парк) заменяются естественными условиями и формами пространственной памяти, которые задействуются в реальной жизни. Совершение покупок в супермаркете, к примеру, несомненно, привычная задача для многих людей, и именно она легла в основу исследования Кирасик (Kirasic, 1991). Испытуемыми Кирасик

были молодые (средний возраст 24 года) и пожилые (средний возраст 60 лет) женщины, а ее экспериментальные задания были связаны с различными формами пространственного интеллекта, которые можно реализовать в бакалейном отделе магазина. Одно из заданий заключалось в том, что испытуемых просили составить оптимальный маршрут для покупки товаров семи видов. Другое задание заключалось в том, что испытуемых просили оценить расстояние, которое необходимо было бы пройти, чтобы взять набор товаров, изображенных на картинке. Все задания выполнялись в двух типах условий: в знакомом для испытуемой супермаркете, в котором она обычно делает покупки, и в не знакомом для испытуемой супермаркете, на изучение которого отводилось 15 минут.

На основании полученных данных было сделано несколько выводов. Молодые испытуемые справлялись чуть лучше пожилых, хотя различие было незначительным и проявлялось при выполнении лишь некоторых заданий. Успешность выполнения заданий в знакомых условиях была выше, хотя опять же различия были незначительными. Большой интерес, чем эти главные эффекты, представляет взаимодействие возраста и экспериментальных условий; привычный контекст пожилым женщинам помогал больше, чем молодым. Рисунок 13.5 иллюстрирует эту закономерность в задании с составлением маршрута и в задании с оценкой расстояния. Хотя то, что пожилым людям знакомая обстановка помогает в большей мере, чем молодым, подтверждается не во всех исследованиях (в действительности, данный феномен имел место при выполнении только двух из четырех заданий, предлагавшихся Кирасик), это довольно типичный результат. А обобщение, которое можно сделать на основе этого исследования, считается вполне доказанным: пожилые люди великолепно справляются с тестами памяти, если при этом оценивается естественная для них деятельность в естественных условиях. Среди других условий, используемых при проведении таких исследований — здания или помещения, в которых работают испытуемые (Uttl & Graf, 1993), знакомые городские кварталы (Rabbitt, 1989) и собственные дома испытуемых (West & Walton, 1985).

Во втором исследовании повседневной памяти охватывался более широкий диапазон ситуаций. В последние годы проводилась разработка ряда шкал памяти, основанных на оценке испытуемыми возможностей собственной памяти в разнообразных контекстах. В табл. 13.5 представлены некоторые вопросы одной из подобных методик: «Клинической шкалы самооценки памяти» (*Memory Assessment Clinics Self-Rating Scale*; Crook & Larrabee, 1990; Winterling, Crook, Salama, & Gobert, 1986), или *MAC-S*. Помимо специфических вопросов, приведенных в таблице, *MAC-S* содержит несколько вопросов общего характера, к примеру: «Если сравнить вашу память в тот период, когда она была лучше всего, сколько времени вам бы потребовалось сегодня, чтобы что-то запомнить или вспомнить?» (от *значительно больше* до *значительно меньше*). Такие методики могут использоваться для исследования ряда вопросов, два из которых мы рассмотрим (более подробно см.: Dixon, 1989; Gilewski & Zelinski, 1986). Первый вопрос касается, конечно же, возрастных изменений; как правило, в исследованиях выявляются проблемы с памятью, которые с возрастом значительно обостряются. Второй вопрос касается соответствия между самоотчетами и реальными возможностями памяти — адекватно ли оценивают люди сильные и слабые стороны своей памяти? Как и в случае с любой мето-

дикой самооценки, ответ на этот вопрос таков — «не совсем»; тем не менее результаты использования большинства таких методик положительно, хотя и не абсолютно, коррелируют с экспериментальными или клиническими оценками памяти. Нужно отметить, что представления людей о собственной памяти интересны сами по себе, даже если они не соответствуют действительности; тот факт, что самооценка обладает прогностической силой, делает ее еще более интересной.

Таблица 13.5

Примеры вопросов из Клинической шкалы самооценки памяти (Memory Assessment Clinics Self-Rating Scale)

Шкала способностей	
Измеряемый параметр	Примеры
Способность помнить специфическую информацию — оценивается по 5-уровневой шкале от <i>очень плохо</i> до <i>очень хорошо</i>	Подробности событий, происходивших в семье за последний год. Имя только что представленного человека. Куда вы кладете свои вещи (например, ключи) дома или на работе. Значения слов, которые вы когда-то неплохо знали. Как добраться до географического пункта, в котором вы были один или два раза
Шкала частоты соответствующих эпизодов	
Измеряемый параметр	Примеры
Частота эпизодов затруднений с памятью на специфическую информацию — оценивается по 5-уровневой шкале от <i>очень часто</i> до <i>очень редко</i>	Трудно вспомнить слово, которые вы хотели употребить. Теряете нить в разговоре. Заходите в комнату, чтобы что-то взять, и забываете, за чем вы туда пошли. Забываете о назначенной встрече или другом событии, которое для вас важно. Не можете узнать человека, который вас узнал

Источник: A Self-Rating Scale for Evaluating Memory in Everyday Life, T. H. Crook and G. J. Larrabee, 1990, *Psychology and Aging*, 5, p. 55.

Рассмотрим еще один вид памяти. Начиная этот раздел, я говорил о стереотипе, согласно которому в пожилом возрасте память ухудшается. Однако этот стереотип требует существенного уточнения: долговременная память может оставаться сохранной до глубокой старости даже при значительном ухудшении памяти на недавние события. Наглядный пример: пожилой человек ярко и в подробностях описывает собственное детство, однако не в состоянии вспомнить, что он ел сегодня на завтрак. Таким образом, ухудшение памяти избирательно (по крайней мере это следует из приведенного выше довода) и эта избирательность удивительна.

Насколько соответствует действительности предположение о сохранности **долговременной памяти**? Вероятно, в определенной степени соответствует, однако,

как оказалось, этот вопрос трудно исследовать, и большая часть данных представляет собой частные случаи. Ниже мы кратко рассмотрим несколько примеров исследований долговременной памяти, а также некоторые методологические трудности, возникающие при проведении подобных исследований (см. также Erber, 1981).

Барик, Барик и Уиттлингер (Bahrick, Bahrick & Wittlinger, 1975) тестировали лиц в возрасте от 17 до 74 лет, закончивших школу. Окончание школы при отборе испытуемых являлось очень важным критерием, поскольку осуществлялась проверка памяти на имена одноклассников. Использовалось 6 тестов памяти: 2 на воспроизведение и 4 на узнавание. Для выполнения тестов на воспроизведение требовалось вспомнить имена как можно большего числа одноклассников, а также указать имена одноклассников, изображенных на фотографиях. Для выполнения тестов на узнавание требовалось выбрать из 5 фотографий фотографию одноклассника, подобрать к имени одну или несколько фотокарточек и подобрать к фотографии одно или несколько имен. Естественно, время хранения информации (то есть время с момента последней встречи с одноклассниками) изменялось с возрастом испытуемых: от 3–4 месяцев у молодых испытуемых до более 50 лет у пожилых.

Не вызывает удивления тот факт, что в этом исследовании молодым испытуемым удавалось вспомнить больше одноклассников, чем пожилым. Возрастные различия были особенно велики в тестах на воспроизведение. В первом задании, к примеру, самые молодые испытуемые в среднем перечислили более 50 имен, а в старшей группе средний уровень воспроизведения составлял около 20 имен. Результаты выполнения тестов на узнавание, напротив, были почти одинаково высокими во всех возрастных группах. В частности, испытуемые, окончившие школу 30–40 лет назад, хорошо узнавали одноклассников по лицам. Успешность выполнения заданий на узнавание согласовывается с приведенными ранее данными о сохранности способности к узнаванию. Более серьезное ухудшение способности к воспроизведению в процессе старения согласуется с результатами исследований, демонстрирующих взаимодействие между возрастом и типом памяти. А то, что определенная информация может удерживаться в памяти в течение столь длительных периодов времени, подтверждает гипотезу о возможности сохранности долговременной памяти до глубокой старости.

Описанное выше исследование — лишь одно из ряда работ, направленных на изучение долговременной памяти. Исследовалась также память на названия улиц и на адреса (Bahrick, 1983), а также память на испанский язык, который испытуемые изучали в школе или колледже (Bahrick, 1984). Другие исследовательские группы анализировали память на имена учителей (Schonfield, 1969), память на названия песен (Barlett & Snelus, 1980) и память на содержание университетских курсов (Cohen, Shanhope, & Conway, 1992). Как и в исследовании с именами и фотографиями одноклассников, в этих исследованиях обычно обнаруживалось, что информация может храниться в памяти удивительно долго и что с возрастом наблюдается некоторое, но всегда существенное, ухудшение памяти.

Наиболее очевидный недостаток подобных исследований, по крайней мере в том, что касается старения, мы уже обсуждали: длительность хранения информации изменяется совместно с возрастом испытуемого. Так, обнаружив, что пожилые испытуемые выполняют тесты памяти не так успешно, как молодые, мы не можем сказать, объясняется это межвозрастными различиями или тем, что удерживать в

памяти информацию в течение 50 лет труднее, чем в течение 5 лет. Скорее всего, имеют значение оба фактора; однако в подобных исследованиях разграничить их невозможно.

Альтернативный подход к изучению долговременной памяти иллюстрируют некоторые другие исследования (Botwinick & Storandt, 1974). Взрослые испытуемые в возрасте от 20 до 79 лет выполняли задания тестов, включавших ряд вопросов об общественно-исторических фактах. Примеры этих вопросов приведены во вставке 13.2. (Для тех, кто хочет себя проверить, ответы представлены в приложении, в конце этой главы.) Поскольку вопросы касаются событий, происшедших на протяжении четырех исторических периодов, появилась возможность рассмотреть память как функцию от возраста испытуемого, давности события, о котором задается вопрос, и возраста испытуемого на момент, когда это событие произошло. Ботвиник и Сторандт в целом не обнаружили никаких возрастных различий в успешности выполнения заданий. То есть 70-летние испытуемые справлялись с заданиями почти так же хорошо, как 20-летние. Однако наблюдалось интересное взаимодействие между возрастом испытуемых и историческим периодом, к которому относился вопрос. Это взаимодействие отражало тот факт, что испытуемые всех возрастов лучше всего помнили о событиях, происшедших в то время, когда им было 15–25-лет. Так, испытуемые в возрасте около 50 лет лучше всего отвечали на вопросы, касавшиеся событий 1930–1940-х годов; испытуемые в возрасте около 30 лет — на вопросы, касавшиеся событий 1950–1960-х годов. Следует добавить, что взаимодействие возраста и исторического периода, к которому относятся вопросы, в подобных исследованиях обнаруживается не всегда (в вышеописанном исследовании оно наблюдалось только у испытуемых мужского пола). Однако результаты других исследований (например, Howes & Katz, 1988; Perlmutter, Metzger, Miller, & Nezworski, 1980) подтверждают главный вывод, к которому пришли Ботвиник и Сторандт: память на общественно-исторические события с возрастом если и ухудшается, то незначительно.

Вышеописанный метод устраняет некоторые недостатки исследований Барик, однако имеет и свои собственные. Если стимулами являются имена и лица одноклассников, можно быть уверенными в том, что испытуемому они когда-то были известны, поэтому любые ошибки при выполнении тестов на воспроизведение или на узнавание могут объясняться неспособностью вспомнить, а не изначальным отсутствием знаний. При использовании вопросов об общественно-исторических фактах вероятность изначального одинакового объема знаний испытуемых меньше. Возможно, некоторые испытуемые никогда и не знали настоящего имени «красного барона». И тогда неспособность дать правильный ответ вряд ли можно расценивать как неспособность вспомнить. Кроме того, спорным моментом является использование термина «память» в ситуации, когда события, о которых спрашивают испытуемого, произошли до его рождения. Вопросы, предлагаемые Ботвиник и Сторандт, не затрагивают специфических, личных воспоминаний, на оценку которых направлено исследование Барик с коллегами. Они направлены на оценку общих знаний, которые могут быть накоплены разными способами.

Каковы бы ни были недостатки исследования Ботвиник и Сторандт, подобные ему исследования иллюстрируют ряд важнейших особенностей памяти взрослых.

Вставка 13.2

Вопросы из теста долговременной памяти Ботвиник и Сторандт**Период: 1950–1960**

Как звали первого человека, ступившего на Луну?

Как звали человека, убившего Мартина Лютера Кинга?

В каком городе был убит президент Джон Ф. Кеннеди?

В каком году русские запустили первый спутник?

Как звали человека, боровшегося за президентское кресло с Дуайтом Эйзенхауэром в 1952 и в 1956 годах?

Как звали сенатора из Висконсина, с именем которого связаны расследования в Конгрессе по делам о коммунистах в 1950-х годах?

Период: 1930–1949

Как звали человека, избранного вице-президентом в 1948 году, когда президентом стал Гарри Трумэн?

Как звали немецкого генерала времен Второй мировой войны, которого называли «лисом пустыни»?

Как звали командира известных «Летающих тигров» времен Второй мировой войны?

Когда (день, месяц, год) японцы осуществили атаку на Перл-Харбор?

Как звали единственного президента США, избранного на этот пост четыре раза?

Где (в каком штате) сгорел и разбился немецкий дирижабль «Фон Гинденбург»?

Период: 1910–1929

Как называлось судно, столкнувшееся с айсбергом и затонувшее во время первого своего плавания в 1912 году?

Как звали человека, смерть которого послужила поводом для развязывания Первой мировой войны?

Как на самом деле звали немецкого летчика-аса времен Первой мировой войны, которого называли «красным бароном»?

Как расшифровывается аббревиатура *ИСТУ*?

Как звали человека, осужденного в известном «обезьяньем процессе» 1925 года?

На каком самолете Линдберг перелетел через Атлантику?

Период: 1890–1909

Как звали человека, открывшего Северный полюс?

В каком штате США впервые была осуществлена казнь на электрическом стуле?

Как звали человека, ставшего президентом после того, как убили президента Маккинли?

Где прошел первый удачный полет братьев Райт?

Как звали боксера, которого все называли «Джентльмен Джим»?

В каком году Генри Форд выпустил Модель Т?

Источник: *Memory, Related Functions, and Age* (p. 189–191), J. Botwinick and M. Storandt, 1974, Springfield, IL: Charles C. Thomas.

В целом они еще раз подтверждают тот факт, что у пожилых людей не обязательно должно наступать значительное ухудшение памяти. С некоторыми заданиями пожилые люди справляются не хуже, а иногда и лучше молодых. Результаты исследования, свидетельствующие о длительном сохранении в памяти информации общего плана, согласуются с результатами других исследований по этому вопросу.

Например, с работами по кристаллизованному интеллекту, описанными ранее. Наконец, исследования, подобные исследованию Ботвиник и Сторандт, указывают на возможность позитивной оценки памяти в старости, если речь идет об осмысленном материале из области общих знаний о мире, а не о произвольных, необычных лабораторных заданиях.

Личность и социальное развитие

Перейдем теперь от интеллекта к личности и социальному аспекту развития. Начнем с рассмотрения двух подходов к изучению личности в период взрослости. Затем мы обсудим некоторые общие вопросы, касающиеся этих и аналогичных подходов.

Удовлетворенность жизнью

Одну из популярных тем психологии старения называют по-разному: «моральное состояние», «субъективное благополучие», «удовлетворенность жизнью» или просто «счастье». Речь идет о вопросе, имеющем огромное значение: насколько доволен человек своей жизнью? Для оценки удовлетворенности жизнью было разработано множество методик. Мы рассмотрим две, начав с одной из первых и наиболее известных — с «Индекса удовлетворенности жизнью» (*Life Satisfaction Index*), или *LSI*, который разработали Нойгартен, Хэвигхерст и Тобин (Neugarten, Havighurst, & Tobin, 1961).

LSI был создан в ходе одного из первых широкомасштабных исследований старения — Канзасского исследования жизни взрослых (*Kansas City Study of Adult Life*). Выборка состояла из 177 мужчин и женщин в возрасте от 50 до 90 лет. С каждым испытуемым в течение 2-летнего периода проводилась серия подробнейших интервью. Как следует из приведенного ниже описания, в интервью анализировался целый ряд аспектов жизни в зрелом и пожилом возрасте.

«Включалась информация о повседневной деятельности; о времяпрепровождении в выходные дни; о членах семьи, с которыми проживают испытуемые; о родственниках, друзьях, соседях; о доходе и о работе; о членстве в общественных организациях; об уровне социальной активности по сравнению с социальной активностью в 45-летнем возрасте; об отношении к старости, болезни, смерти и бессмертию; об одиночестве, скуке, гневе; а также о лицах, которые испытуемые воспринимают как образцы для подражания, и об образе собственного Я» (Neugarten et al., 1961, p. 136).

Задача интервью состояла в получении информации, позволяющей выделить различные аспекты удовлетворенности жизнью, подлежащие количественной оценке. В конечном счете было выделено пять параметров: «Энергичность/Апатия», «Решительность и стойкость», «Конгруэнтность желаемого и достигнутого», «Я-концепция» и «Настроение». Каждый параметр оценивался по 5-уровневой шкале, причем 5 соответствовало положительному полюсу, а 1 — отрицательному.

Метод интервью, хотя и необходим для предварительного очерчивания границ каждой из сфер, все же требует слишком больших затрат времени и усилий на оценку каждого испытуемого. Поэтому следующим этапом исследования стала разра-

ботка краткой методики для получения той же информации. Результатом явился *LSI* — опросник, основанный на самоотчете, при заполнении которого от испытуемого требуется только согласиться или не согласиться с 20 утверждениями. В табл. 13.6 представлены 20 утверждений опросника, а также ответы, которые дали бы максимальный показатель удовлетворенности по всем параметрам. Как показывают исследования, коэффициент корреляции между показателями *LSI* и рейтинговыми оценками, полученными при использовании метода интервью, составляет 0,55 — то есть корреляционная связь существует, но далеко не абсолютная.

Как отмечалось, *LSI* направлен на измерение удовлетворенности жизнью в пожилом возрасте. Однако разумеется, удовлетворенность жизнью — проблема, актуальная не только на последних этапах жизненного пути. Более того, любая попытка оценить стабильность или изменчивость удовлетворенности жизнью в период взрослости предполагает использование методики, подходящей для разных возрастных групп. Одна из методик широкого применения — «Шкала удовлетворенности жизнью (*Satisfaction with Life Scale*), или *SWLS* (Diener, Emmons, Larsen, & Griffin, 1985). *SWLS* невелика по объему и поэтому удобна в использовании; все пять утверждений, составляющие этот тест, представлены в табл. 13.7. Несмотря на краткость *SWLS*, ее валидность как методики измерения субъективного благополучия на разных этапах периода взрослости подтверждается разнообразными данными (Myers & Diener, 1995; Pavot, Diener, Colvin, & Sandvik, 1991). Было доказано, к примеру, что результаты этого теста коррелируют с результатами более длинных методик оценки удовлетворенности жизнью. Они также согласуются с информацией о моральном состоянии человека, полученной из других источников, например, из отчетов друзей или из клинических оценок.

Что же удалось узнать об удовлетворенности жизнью в пожилом возрасте с помощью таких методик, как *LSI* и *SWLS*? Зададим сначала важнейший для психологии развития вопрос: изменяется ли удовлетворенность жизнью с возрастом? Ответить на этот вопрос сложно, поскольку большинство исследований по этой теме проводилось методом поперечных срезов, поэтому выявленные возрастные изменения могут в действительности отражать различия между когортами, а не истинную динамику. В любом случае, результаты противоречивы. В одних исследованиях обнаруживается в целом отрицательная связь между возрастом и удовлетворенностью жизнью, по крайней мере у лиц в возрасте 60 лет и старше. Эдвардс и Клеммак (Edwards & Klemmack, 1973), к примеру, сообщают о коэффициенте корреляции между возрастом и показателями *LSI* (в выборке из 507 испытуемых среднего и пожилого возраста), равном $-0,14$. С другой стороны, во многих исследованиях такой связи выявить не удалось; в действительности, это, наверное, самый типичный вывод, полученный в последних исследованиях (Myers & Diener, 1995). И даже когда некие эффекты все же выявляются, связь между возрастом и удовлетворенностью жизнью оказывается в лучшем случае умеренной. То есть опять-таки следует отметить наличие существенных индивидуальных различий в пожилом возрасте. Многие люди преклонного возраста вполне довольны собственной жизнью.

Индекс удовлетворенности жизнью

Ниже представлены утверждения о жизни в целом, в отношении которых люди имеют разные мнения. Прочтите каждое утверждение списка и, если вы с ним согласны, поставьте крестик напротив него в графе «Согласен». Если вы не согласны с утверждением, поставьте напротив него крестик в графе «Не согласен». Если вы не уверены, поставьте крестик напротив утверждения в графе «?» **ПРОВЕРЬТЕ, ПОЖАЛУЙСТА, ОТВЕТИЛИ ЛИ ВЫ НА ВСЕ ВОПРОСЫ СПИСКА.**

(Ключ: 1 балл за каждый крестик)

	Со- гла- сен	Не согла- сен	?
1. Становясь старше, я обнаруживаю, что все не так плохо, как я ожидал			
2. В моей жизни было больше неудач, чем в жизни большинства тех, кого я знаю			
3. Это самое ужасное время в моей жизни			
4. Сегодня я чувствую себя не менее счастливым, чем в молодости			
5. Моя жизнь могла бы быть лучше, чем она есть на самом деле			
6. Это лучшие годы моей жизни			
7. Я занимаюсь в основном скучными и монотонными делами			
8. От будущего я ожидаю приятных и интересных событий			
9. То, что я делаю сегодня, интересно мне не меньше, чем то, что я делал раньше			
10. Я чувствую себя старым и уставшим			
11. Я чувствую свой возраст, но это меня не беспокоит			
12. Оглядываясь на свое прошлое, я доволен прожитой жизнью			
13. Я бы не изменил своего прошлого, даже если бы мог			
14. По сравнению с другими людьми моего возраста я наделал в своей жизни гораздо больше глупостей			
15. По сравнению с другими людьми моего возраста я выгляжу хорошо			
16. У меня есть планы на месяц или на год вперед			
17. Когда я размышляю о прожитой жизни, то понимаю, что не сделал большую часть того, что хотел сделать			
18. По сравнению с другими людьми я впадаю в уныние слишком часто			
19. Я ожидаю многого от жизни			
20. Что бы ни говорили люди, жизнь среднего человека становится печальнее, а не радостнее			

Источник: The Measurement of Life Satisfaction, B. L. Neugarten, R. J. Havinghurst, and S. S. Tobin, 1961, *Journal of Gerontology*, 16, p. 141.

Таблица 13.7

Шкала удовлетворенности жизнью

	Вопросы	Инструкция
1	В основном моя жизнь соответствует моему идеалу	«Ниже представлены пять утверждений, с которыми вы можете согласиться или не согласиться. Используя шкалу от 1 до 7, укажите, насколько вы согласны с этими утверждениями, поставив соответствующую цифру перед каждым утверждением. Пожалуйста, в своих ответах будьте открыты и честны». Шкала: от 1 (<i>не согласен</i>) до 7 (<i>полностью согласен</i>)
2	Обстоятельства моей жизни складываются благополучно	
3	Я удовлетворен своей жизнью	
4	Пока у меня есть все, что мне нужно от жизни	
5	Если бы я мог прожить свою жизнь вновь, я бы почти ничего не изменил	

Источник: The Satisfaction with Life Scale, E. Diener, R. A. Emmons, R. J. Larsen, & S. Griffin, 1985, *Journal of Personality Assessment*, 49, p. 72.

Представленные выше результаты естественным образом подводят нас к вопросу о том, каковы детерминанты возрастных и индивидуальных различий в удовлетворенности жизнью. На разрешение этой проблемы было направлено множество исследований с использованием *LSI* и *SWLS* и аналогичных методик. Был выявлен ряд коррелятов, многие из которых обладают прогностической силой (Diener, 1984; Myers & Diener, 1995). С удовлетворенностью жизнью, к примеру, связано состояние здоровья; в целом здоровые люди больше довольны своей жизнью, чем люди не очень здоровые. Связь с удовлетворенностью жизнью имеет и семейное положение; у лиц, состоящих в браке, показатели удовлетворенности жизнью в целом выше, чем у овдовевших или разведенных. Удовлетворенность жизнью также имеет связь с доходом, причем в ожидаемом направлении. Наконец, с удовлетворенностью жизнью связаны переменные поведенческого и психического плана. Наверное, наиболее интересный результат касается связи между уровнем активности и удовлетворенностью жизнью. В ряде исследований было обнаружено, что эта связь положительна — то есть пожилые люди, ведущие активный образ жизни, более довольны своей жизнью, чем те, у кого уровень активности снижен. Вскоре мы вернемся к этому вопросу.

Методики изучения личностных особенностей

Каким бы глобальным и важным ни было понятие удовлетворенности жизнью, оно определяет лишь один из аспектов личности взрослого человека. Методики, к рассмотрению которых мы сейчас приступим, направлены на измерение характеристик значительно более широкого диапазона, так как назначение этих методик — всесторонняя оценка личности и индивидуальных различий. Нет необходимости говорить, насколько по-разному подходят различные специалисты к определению понятия личности и насколько много поэтому методик ее оценки. Олдвин и Левенсон (Aldwin & Levenson, 1994) провели заслуживающий внимания обзор некоторых методик, наиболее часто используемых при изучении старения. Пример, на

котором мы остановимся, — главная методика, использовавшаяся в одной широкомасштабной исследовательской программе, — «Личностный опросник *NEO*» (*NEO Personality Inventory*), или *NEO PI-R* (Costa & McCrae, 1992).

NEO PI-R может использоваться и как методика самоотчета, и как методика отчета о третьем лице. В обоих случаях он состоит из 240 утверждений, описывающих различные аспекты личности взрослого. В табл. 13.8 приводятся примеры утверждений, сформулированные в версии самоотчета. Испытуемый оценивает каждое из утверждений по 5-уровневой шкале: от *совершенно не согласен* до *полностью согласен*.

Уже из этих примеров должно быть видно, что с помощью *NEO PI-R* можно получить значительный объем разнообразной информации о личности. Но что делают с этой информацией потом, то есть как на основе 240 индивидуальных утверждений делают выводы о личностных особенностях? Решая этот вопрос, Коста и Маккрей — как и создатели множества других тестов личности — использовали такую статистическую процедуру, как факторный анализ. Назначение факторного анализа — выделение ряда компонентов, или «факторов», которые можно выявить при измерении некоего конструкта. Этот метод (значительно более сложный, чем представленное здесь описание) связан с анализом паттернов корреляции в совокупности оценочных суждений; при этом предполагается, что суждения, тесно связанные друг с другом корреляционными связями, измеряют один и тот же фактор. Используя результаты предыдущих исследований, теорию и факторный анализ собственной тестовой батареи, Коста и Маккрей выделили 5 факторов, объясняющих большую часть индивидуально-личностных различий. Первые 3 из этих факторов определяют аббревиатуру *NEO*: нейротизм, экстраверсия и открытость (*Neuroticism, Extraversion, Openness*). Два другие фактора были добавлены позже: это уступчивость (*Agreeableness*) и добросовестность (*Conscientiousness*). В табл. 13.8 приведены некоторые из утверждений, направленные на оценку каждого из факторов. (Обратите внимание, что в форме, предлагаемом испытуемому, вопросы не сгруппированы по факторам, как в таблице, а распределены по всему опроснику.)

Как всегда, создание методики — лишь первый шаг исследования. *NEO* и другие личностные опросники используются при изучении разнообразных вопросов. Здесь я сосредоточусь на одном из важнейших вопросов, возникающих при исследовании старения, — это проблема стабильности личностных черт. Можно ли предположить, что конstellация характеристик, определяющих личность в 20-летнем возрасте, останется неизменной и в 40, и в 60, и в 80 лет? Или личностные особенности с возрастом изменяются? И отличаются ли в среднем личностные характеристики пожилых людей от личностных характеристик молодых?

Основным источником информации по этой теме служат результаты использования *NEO PI-R* и методик, созданных ранее (Costa & McCrae, 1989; McCrae & Costa, 1990, 1994). В исследованиях проводилось как поперечное сравнение испытуемых разных возрастов, так и лонгитюдное исследование тех же испытуемых. Данные обоих типов убедительно свидетельствовали о стабильности личностных черт. При использовании метода поперечных срезов были выявлены определенные

Таблица 13.8

Примеры утверждений из личностного опросника NEO (NEO Personality Inventory)

Параметр	Утверждения
Нейротизм	Я часто испытываю напряжение и нервничаю Я часто сержусь на то, как люди ведут себя по отношению ко мне Я редко чувствую себя одиноким и редко впадаю в уныние Мне часто бывает трудно принять решение
Экстраверсия	Мне нравится большинство людей, с которыми я встречаюсь Я стараюсь избегать большого скопления людей Я часто оказываюсь душой компании Если я что-то делаю, то делаю это с душой
Открытость	У меня сильно развито воображение Мои эмоции и чувства разнообразны Я никогда не изменяю своим привычкам Мне нравится рассуждать на отвлеченные темы
Уступчивость	Я уверен, что большинство людей действуют из лучших побуждений Иногда я манипулирую людьми, чтобы добиться желаемого Я предпочитаю сотрудничать с другими людьми, а не соревноваться с ними Я не люблю говорить о себе и своих достижениях
Добросовестность	Я бываю требовательным и привередливым Я стараюсь выполнять порученные мне задания добросовестно Мне трудно заставить себя сделать то, что я должен сделать Я редко принимаю поспешные решения

Источник: *Manual for the Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI)* (p. 68–74), P. T. Costa, Jr., and R. R. McCrae, 1992. Odessa, FL: Psychological Assessment Resources, Inc.

изменения в уровне выраженности черт до 30-летнего возраста, однако никаких свидетельств систематических изменений после этого возраста получено не было. Так, 70-летние испытуемые за редкими исключениями демонстрировали те же паттерны ответов по NEO, что и 30-летние. При использовании лонгитюдного метода были получены данные, указывающие на заметную стабильность личностных характеристик на протяжении периода взрослости при коэффициенте корреляции от 0,6 до 0,8 даже с временным интервалом 30 лет. Коста и Маккрей пишут: «Стабильность, по-видимому, характерна для всех пяти важнейших аспектов личности... [Данные] свидетельствуют о том, что профиль личности взрослого человека в целом практически не изменяется» (McCrae & Costa, 1994, p. 173).

Акцентируя внимание на этой исследовательской программе, я должен заметить, что не все придерживаются точки зрения авторов на проблему стабильности (Aldwin & Levenson, 1994; Helson, 1993). Данные, которые получили Коста и Маккрей, могут интерпретироваться по-разному, и есть основания полагать, что при

использовании других подходов к оценке личности были бы выявлены более значительные изменения, чем при оценке отдельных черт с помощью *NEO PI-R*. Тем не менее не будет ошибкой сказать, что их работа убедила почти всех, кто изучает старение, в том, что личность значительно стабильнее, чем считалось ранее. Результаты этой работы также позволяют опровергнуть еще один стереотип, касающийся старения, — стереотип личности пожилого человека. В старости некоторые люди действительно становятся более ригидными, более замкнутыми, подавленными и т. п., однако это — индивидуальные изменения, являющиеся скорее исключением, чем правилом.

Общие замечания

Обратимся теперь к некоторым общим методологическим вопросам, имеющим отношение к личностным опросникам, которые мы привели в качестве примеров. Первое, о чем следует упомянуть, — это только некоторые примеры, малая часть методик, которые можно использовать при изучении периода взрослости и старения. «Личность», как и «интеллект», нельзя свести к одному тесту, и даже такие специфические конструкты, как «удовлетворенность жизнью» могут истолковываться и оцениваться по-разному. Хотя результаты использования различных методик, естественно, во многом сходятся, это сходство далеко не полное. Таким образом, то, к какому выводу мы придем, будет зависеть от избранной нами методики.

Второй момент касается вопроса об эквивалентности измерения. Как неоднократно подчеркивалось, выводы о возрастных различиях или возрастных изменениях валидны лишь тогда, когда используемые методики одинаково подходят для изучения сравниваемых групп. Как и тест интеллекта, многие личностные опросники, используемые при работе с пожилыми людьми, изначально разрабатывались для лиц молодого и среднего возраста. Даже методики, при создании которых охватывался весь жизненный путь человека, по содержанию ближе молодым испытуемым. И даже когда содержание соответствует возрасту, на ответы пожилых людей могут влиять факторы, связанные с процессом тестирования (например, трудности при заполнении длинных опросников).

Третий вопрос касается общего подхода к оценке личностных особенностей. В этой книге рассматриваются три основных подхода к изучению любой сферы поведения: непосредственное наблюдение поведения, интересующего исследователя, в естественных условиях, экспериментальное стимулирование поведения в структурированных лабораторных условиях и рейтинг, или вербальный отчет о типичном поведении, получаемый от тех, кто знает испытуемого. Методики, которые рассматривались в этом разделе, однозначно можно отнести к третьей категории — то есть к рейтингу, или вербальным отчетам. До сегодняшнего времени при оценке личностных особенностей взрослых, как правило, используется метод рейтинга. Однако существуют и исключения. Для изучения определенной проблематики иногда применяется экспериментальный подход; примеры подобных тем — конформность (например, Klein, 1972), ригидность (например, Ohta, 1981) и осторожность (например, Okun & Elias, 1977). Существуют, хотя и немногочисленные, данные об использовании обсервативной оценки — к примеру, обсервативной оценки пожилых людей, проживающих в домах престарелых (Martino-Salzman, Blasch,

Morris, & McNeale, 1991). Тем не менее основная часть информации о личности и старении, которую мы имеем, была получена с использованием метода рейтинга. Поэтому при оценке литературы по этой тематике, следует постоянно помнить о плюсах и минусах этого подхода, которые обсуждались ранее.

Стоит отметить еще один момент. Упомянутые выше методики основаны не просто на вербальных отчетах; в большинстве случаев это *самоотчеты*. Информация об испытуемом предоставляется самим испытуемым. Поэтому все затруднения и искажения, сопутствующие использованию методик самоотчета, имеют отношение и к этим тестам. Лоутон, Уэлинан и Белски (Lawton, Whelinan, & Belsky, 1980) рассмотрели в своей работе факторы искажения, возможные в процессе тестирования, а также способы минимизации этих искажений при оценке личностных особенностей пожилых испытуемых. Исследователи указывают, что подобные искажения (например, эффект ожидания оценки, установка на определенный ответ, неправильное понимание инструкций, утомление) могут влиять не только на выводы об особенностях пожилого возраста, но и на выводы о личностно-возрастных различиях. Кажущиеся различия в уровне депрессии, к примеру, могут отражать межвозрастные различия в готовности признаться в собственных негативных переживаниях, а не истинные личностные различия между молодыми и пожилыми людьми.

Последний вопрос касается корреляционного характера большей части рассмотренных данных. Представляющая интерес связь между личностью и возрастом, конечно же, корреляционная. Кроме того, во многих из упомянутых выше исследовательских программ основной упор делался на установление корреляций. Возьмем, к примеру, вопрос о детерминантах удовлетворенности жизнью. Играла ли роль состояние здоровья? Оказывает ли влияние уровень активности? Ответ зависит от величины коэффициента корреляции. Нас интересуют причинно-следственные связи (например, активность как причина удовлетворенности жизнью), а работаем мы с неэкспериментальными исследовательскими планами, не позволяющими выявить причинность. Любые корреляционные связи имеют несколько возможных объяснений. В случае с активностью и удовлетворенностью жизнью вполне вероятно, что удовлетворенность жизнью обусловлена активным образом жизни. Но вероятно также, что направление причинного воздействия обратное: удовлетворенность жизнью помогает людям вести активный образ жизни. И несомненно, что оба фактора влияют друг на друга: активность способствует удовлетворенности, а удовлетворенность поддерживает активность.

Недостатки корреляционных исследований и способы их устранения рассматривались в главе 3. Как мы убедились, на примере многих исследований не стоит сетовать на корреляционный характер данных, поскольку иногда выявить корреляционные связи — это лучшее, что можно сделать. Возьмем, к примеру, вывод о том, что семейное положение связано с удовлетворенностью жизнью. Провести эксперимент, спровоцировав развод пожилых пар с тем, чтобы определить его влияние на дальнейшую жизнь испытуемых, не представляется возможным. Однако что-то для выявления причинно-следственных связей мы все же можем сделать. Во-первых, использовать различные статистические приемы (например, метод парциальных корреляций, описанный в главе 3), предназначенные для контроля

потенциального смещения факторов и более точного определения связи между соответствующими переменными. Можно, к примеру, задаться вопросом, как изменится связь между семейным положением и удовлетворенностью жизнью, если мы будем контролировать социально-экономический статус. Такие статистические приемы контроля часто используются при изучении рассмотренных выше тем. Еще один способ — изучать паттерны корреляций во времени, основываясь на том, что причина неизменно предшествует следствию. Например, снижение удовлетворенности жизнью после развода — более надежное основание для вывода о причинно-следственной связи, чем констатация корреляции при однократном измерении. Здесь вновь следует отметить огромное значение лонгитюдного метода при изучении ключевых вопросов психологии развития.

Резюме

Первая часть этой главы посвящена общим вопросам изучения старения. В ней также анализируется одна из всесторонне изученных тем — стабильность/изменчивость *IQ*.

При исследовании старения возникают те же методологические трудности, что и в любом исследовании из области психологии развития, однако здесь они становятся наиболее актуальными. При *отборе* преследуются две цели: достижение репрезентативности выборок разных возрастов и достижение сопоставимости этих выборок. Достигнуть и той и другой цели достаточно трудно, кроме того, они могут противоречить друг другу. Сравнение молодых и пожилых испытуемых особенно осложняется возможностью двух типов смещения факторов. Первый — смещение возраста и уровня образования: в среднем молодые люди более образованы, чем пожилые. Второй тип смещения — смещение возраста и состояния здоровья: в среднем у молодых людей состояние здоровья лучше, чем у пожилых. Оба фактора обуславливают снижение показателей *IQ* с возрастом.

Обсуждение переменной состояния здоровья заставляет задуматься о проблемах общего характера, возникающих в лонгитюдных исследованиях. Одна из важнейших — *избирательное выбывание*. В исследованиях интеллекта чаще всего выбывают испытуемые со сравнительно низким *IQ*. Это обстоятельство имеет два основания: выход из исследования по собственному желанию и выход из исследования, не зависящий от воли испытуемого и связанный с ухудшением здоровья или в крайнем случае — смертью. Результатом этого является искажение результатов в сторону улучшения показателей пожилых людей.

Затем обсуждается вопрос *измерения*. Как и любой глобальный конструкт, *IQ* можно измерять разными способами. Чаще всего в целях измерения интеллекта у пожилых людей используют *WAIS* и *РМА*. Оба теста делятся на подшкалы, вероятность межвозрастных различий по которым неодинакова. Возрастные различия по шкалам вербального интеллекта выражены меньше, чем по шкалам невербального интеллекта, по показателям кристаллизованного интеллекта — меньше, чем по показателям текучего интеллекта, по нескоростным методикам — меньше, чем по скоростным. Кроме того, возрастные различия с меньшей вероятностью выявляются с помощью методик измерения так называемого «повседнев-

ного» интеллекта. При обсуждении альтернатив *IQ* встает вопрос, имеющий первостепенное значение при любом сравнении молодых и пожилых людей, — проблема *эквивалентности измерения*. Тесты и контексты, используемые для оценки навыков пожилых людей, часто больше подходят для молодых испытуемых, и этот фактор также может обусловить обнаружение возрастных различий.

Первый раздел главы завершает рассмотрение *плана* исследования. Планы *поперечных срезов* связаны с невозможностью разграничить эффект возраста и когорты; *лонгитюдные* планы связаны с невозможностью разграничить эффект возраста и времени измерения. Значение таких переменных, как когорта и время проведения измерений подтверждается исследованиями с использованием третьего «простого» типа плана — плана с *временным лагом*. Указанные выше смещения наряду с избирательным выбыванием в лонгитюдных исследованиях объясняют тот факт, что в исследованиях методом поперечных срезов потеря навыков с возрастом оказывается значительно больше, чем в лонгитюдных исследованиях. Используя *последовательный* план, можно попытаться устранить недостатки простых планов исследования. Последовательные планы подразумевают сочетание простых планов, которое позволяет выявить роль таких переменных, как возраст, время измерения и когорты. Однако даже последовательные планы имеют свои слабые стороны; поэтому вопрос об относительном значении возраста, когорты и времени измерения как факторов, определяющих успешность выполнения *IQ*-тестов, остается спорным.

Вторая часть главы посвящена специфическим вопросам психологии старения. Одна из популярных тем исследования — память. Как и в случае с *IQ*, вероятность возрастных различий в памяти зависит от ряда факторов. Хотя есть и исключения, но с возрастом способность к *узнаванию* страдает меньше, чем способность к *воспроизведению*, а на выполнении заданий, предполагающих *воспроизведение с подсказкой*, старение отражается в меньшей степени, чем на выполнении обычных тестов памяти. Оба эти обстоятельства наводят на мысль о том, что пожилые люди сталкиваются с особыми трудностями при извлечении информации, хранящейся в памяти. В целом, предполагается, что пожилые люди испытывают затруднения в ситуациях, где необходимо использование *мнемонических стратегий*. Эту гипотезу подтверждают данные исследований, проведенных различными методами. Данные также свидетельствуют о том, что с возрастом ухудшается *конструктивная память*, что, вероятно, обусловлено общим снижением скорости переработки информации. Наконец, более позитивную характеристику памяти у пожилых людей дают результаты исследования двух типов: исследований *«повседневной памяти»* и исследования *долговременной памяти*.

Главу завершает рассмотрение методов изучения личностных особенностей и социальных отношений в пожилом возрасте. В качестве примеров описываются некоторые тесты и приводятся некоторые результаты, полученные с помощью *LSI* и *SWLS* — методик, разработанных для оценки *удовлетворенности жизнью* или субъективного благополучия. *NEO PI-R* — методика, охватывающая более широкий диапазон характеристик; это одна из ряда методик вербального отчета (обычно самоотчета), предназначенных для измерения индивидуальных различий в личностных особенностях у взрослых. Главный вопрос, к которому обращаются исследователи, используя методики обоих типов, — вопрос возрастной стабильности

или изменчивости. Хотя с возрастом изменения, несомненно, происходят, исследования свидетельствуют о том, что многие аспекты личности в период взрослости остаются неизменными.

За рассмотрением конкретных тестов следует обсуждение общих вопросов исследования личностных особенностей пожилых людей. Одна из проблем — проблема эквивалентности измерения: одинаково ли подходят тесты для сравниваемых возрастных групп? Еще одна проблема — проблема валидности: действительно ли эти тесты измеряют то, что они призваны измерять? Особенно сомнительна валидность методик самоотчета, что создает вероятность разнообразных искажений результатов. Последний из рассматриваемых вопросов — корреляционный характер многих исследований старения. С помощью корреляционного анализа невозможно с точностью установить причинно-следственные отношения (например, между активностью и удовлетворенностью жизнью), которые интересуют исследователя.

Упражнения

В исследовании Шайе (Schaie) основное внимание уделяется интеллекту. Предположим, у вас была бы возможность провести аналогичное исследование, направленное на некую другую характеристику. Какую зависимую переменную вы бы выбрали и почему? Какое соотношение влияния возраста, когорты и времени измерения вы бы ожидали обнаружить?

Как отмечалось, среди людей распространены негативные стереотипы старости. Эти стереотипы сами стали объектом исследования; при этом стоял вопрос о том, насколько они укоренились в сознании испытуемых разных возрастов (Awstin, *Gerontologist*, 1985, 25, 431–434; Braithwaite et al., *Australian Psychologist*, 1993, 28, 9–15; Seccombe & Ishii-Kuntz, *Gerontologist*, 1991, 31, 527–533). Выберите какую-либо методику измерения восприятия старости и обследуйте с ее помощью людей разных возрастов. Постарайтесь протестировать по крайней мере одного пожилого испытуемого; если методика это позволяет, было бы интересно протестировать и детей.

В этой главе не раз подчеркивалось значение проблемы эквивалентности измерения. Представьте, что перед вами стоит задача создания методики оценки «повседневного интеллекта», предназначенной для вашей возрастной группы. Какие задания вы бы включили в эту методику, насколько она была бы похожа на методику, описанную в табл. 13.3? Можно ли, по вашему мнению, разработать тест, одинаково валидный для использования и с 20-летними, и с 70-летними испытуемыми?

Приложение

Ответы на вопросы Ботвиник и Сторандт (Botwinick & Storandt, 1974).

1. Нейл Армстронг.
2. Джеймс Ерл Рей.
3. Даллас.
4. 1957.
5. Адлаи Стивенсон (Adlai Stevenson).
6. Джозеф Маккарти.
7. Олбен Баркли (Alben Barkley).
8. Эрвин Роммель.
9. Клэр Шено (Claire Chennault).
10. 7 декабря 1941 года.
11. Франклин Рузвельт.
12. Нью-Джерси.
13. Титаник.
14. Эрцгерцог Фердинанд.
15. Барон Манфред Фон Рихтенховен.
16. Women's Christian Temperance Union (Христианский женский союз воздержания)
17. Джон Т. Скоупс (John T. Scopes).
18. Дух Сент-Льюиса.
19. Роберт Пири (Robert E. Peary).
20. Нью-Йорк.
21. Теодор Рузвельт.
22. Китти Хок (Kitty Hawk).
23. Джеймс Корбетт (James Corbett).
24. 1908.

Литература

Abramovitch, R., Freedman, J. L., Thoden, K., & Nikolich, C. (1991). Children's capacity to consent to participation in psychological research: Empirical findings. *Child Development*, 62, 1100–1109.

Achenbach, T. M. (1978). *Research in developmental psychology: Concepts, strategies, methods*. New York: The Free Press.

Acredolo, L. P., & Hake, J. L. (1982). Infant perception. In B. B. Wolman (Ed.), *Handbook of developmental psychology* (pp. 244–283). Englewood Cliffs, NJ: Prentice-Hall.

Ainsworth, M. D. S. (1967). *Infancy in Uganda: Infant care and growth of love*. Baltimore: Johns Hopkins University Press.

Ainsworth, M. D. S., Bell, S. M., & Stayton, D. J. (1974). Infant mother attachment and social development: «Socialization» as a product of reciprocal responsiveness to signals. In M. P. M. Richards (Ed.), *The integration of a child into a social world* (pp. 99–135). Cambridge, Eng.: Cambridge University Press.

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation*. Ithaca, NY: Erlbaum.

Aldwin, C. M., & Levenson, M. R. (1994). Aging and personality assessment. *Annual Review of Gerontology and Geriatrics*, 14, 182–209/

Als, H., Tronick, E., & Brazelton, T. B. (1979). Analysis of face-to-face interaction in infant-adult dyads. In M. E. Lamb, S. J. Suomi, & G. R. Stephenson (Eds.), *Social interaction analysis: Methodological issues* (pp. 33–76). Madison: The University of Wisconsin Press.

American Psychological Association (1992). Ethical principles of psychologists and code of conduct. *American Psychologist*, 47, 1597–1611.

American Psychological Association. (1994). *Publication manual of the American Psychological Association* (4th ed.). Washington, DC: Author.

Amsterdam, B. (1972). Mirror self-image reactions before age 2. *Developmental Psychobiology*, 5, 297–305.

Anders, T. F., & Roffwarg, H. P. (1973). The effects of selective interruption and deprivation of sleep on the human newborn. *Developmental Psychobiology*, 6, 77–89.

Anstey, K., Stankov, L., & Lord, S. (1993). Primary aging, secondary aging, and intelligence. *Psychology and Aging*, 8, 562–570.

Applebaum, M. I., & McCall, R. B. (1983). Design and analysis in developmental psychology. In P. H. Mussen (Series Ed.) & W. Kessen (Vol. Ed.), *Handbook of child psychology: Vol. 1. History, theory, and methods* (4th ed., pp. 415–476). New York: Wiley.

Arend, R., Gove, F. L., & Sroufe, L. A. (1979). Continuity of individual adaptation from infancy to kindergarten: A predictive study of ego-resiliency and curiosity in preschoolers. *Child Development*, 50, 950–959.

Ashmead, D. H., & Perlmuter, M. (1980). Infant memory in everyday life. In M. Perlmuter (Ed.), *New directions for child development: Vol. 10. Children's memory* (pp. 1–16). San Francisco: Jossey-Bass.

Aslin, R. N., Pisoni, D. B., & Jusczyk, P. W. (1983). Auditory development and speech perception in infancy. In P. H. Mussen (Series Ed.) & M. M. Haith & J. J. Campos (Vol. Eds.), *Handbook of child psychology: Vol. 2. Infancy and developmental psychobiology* (4th ed., pp. 573–687). New York: Wiley.

Astington, J. W. (1993). *The child's discovery of the mind*. Cambridge, MA: Harvard University Press.

Astington, J. W., & Jenkins, J. M. (1995, March). Language and theory of mind: A theoretical review and a longitudinal study. In J. W. Astington (Chair), *Language development and the acquisition of a theory of mind*. Symposium conducted at the meeting of the Society for Research in Child Development, Indianapolis.

Bahrack, H. P. (1983). The cognitive map of a city—50 years of learning and memory. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 17, pp. 125–163). New York: Academic Press.

Bahrack, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General*, 113, 1–26.

Bahrack, H. P., Bahrack, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General*, 104, 54–75.

Baillargeon, R. (1987). Object permanence in 3.5- and 4.5-month-old infants. *Developmental Psychology*, 23, 655–664.

Baillargeon, R. (1993). The object concept revisited: New directions in the investigation of infants' physical knowledge. In C. Granrud (Ed.), *Visual perception and cognition in infancy* (pp. 265–315). Hillsdale, NJ: Erlbaum.

Baillargeon, R. (1994). How do infants learn about the physical world? *Current Directions in Psychological Science*, 3, 133–140.

Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). *Life-span developmental psychology: Introduction to research methods*. Monterey, CA: Brooks/Cole.

Bandura, A. (1989). Social cognitive theory. In R. Vasta (Ed.), *Annals of child development* (Vol. 6, pp. 1–60). Greenwich, CT: JAI Press.

Bandura, A., Grusec, J. E., & Menlove, F. L. (1967). Some social determinants of self-monitoring reinforcement systems. *Journal of Personality and Social Psychology*, 5, 449–455.

Bandura, A., Ross, D., & Ross, S. A. (1963a). Imitation of film-mediated aggressive models. *Journal of Abnormal and Social Psychology*, 66, 3–11.

Bandura, A., Ross, D., & Ross, S. A. (1963b). Vicarious reinforcement and imitative learning. *Journal of Abnormal and Social Psychology*, 67, 601–607.

Barber, T. X. (1976). *Pitfalls in human research: Ten pivotal points*. New York: Pergamon Press.

Barber, T. X., & Silver, M. J. (1968). Fact, fiction, and the experimenter bias effect. *Psychological Bulletin Monograph Supplement*, 70, 1–29.

Barker, R. G., & Wright, H. F. (1951). *One boy's day*. New York: Harper & Row.

Barlow, D. H., & Hersen, M. (1984). *Single case experimental designs*. Elmsford, NY: Pergamon Press.

Barnett, M. A., King, L. M., & Howard, J. A. (1979). Inducing affect about self or other: Effects on generosity in children. *Developmental Psychology*, 15, 164–167.

Barrett, D. E., & Yarrow, M. R. (1977). Prosocial behavior, social inferential ability, and assertiveness in children. *Child Development*, 48, 475–481.

Bartlett, J. C., & Snelus, P. (1980). Lifespan memory for popular songs. *American Journal of Psychology*, 93, 551–560.

Bartsch, K., & Wellman, H. M. (1989). Young children's attribution of action to beliefs and desires. *Child Development*, 60, 946–964.

Bates, J. E., & Wachs, T. D. (1994). *Temperament: Individual differences at the interface of biology and behavior*. Washington, DC: American Psychological Association. Wiley.

Bauer, P. J., & Mandler, J. M. (1989). One thing follows another: Effects of temporal structure on 1- to 2-year-olds' recall of events. *Developmental Psychology*, 25, 197–206.

Bauer, P. J., & Mandler, J. M. (1992). Putting the horse before the cart: The use of temporal order in recall of events by one-year-old children. *Developmental Psychology*, 28, 441–452.

Baumrind, D. (1971). Current patterns of parental authority. *Developmental Psychology Monograph*, 4, 1–103.

Baumrind, D. (1985). Research using intentional deception: Ethical issues revisited. *American Psychologist*, 40, 165–174.

Baumrind, D. (1989). Rearing competent children. In W. Damon (Ed.), *Child development today and tomorrow* (pp. 349–378). San Francisco: Jossey-Bass.

Bayley, N. (1993). *Bayley Scales of Infant Development: Birth to Two Years* (2nd ed.). New York: The Psychological Corporation.

Beal, C. R. (1994). *Boys and girls: The development of gender roles*. New York: McGraw-Hill.

Beere, C. A. (1990). *Gender roles: A handbook of tests and measures*. New York: Greenwood.

Beilin, H. (1965). Learning and operational convergence in logical thought development. *Journal of Experimental Child Psychology*, 2, 317–339.

Bell, R. Q. (1968). A reinterpretation of the direction of effects of socialization. *Psychological Review*, 75, 81–95.

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42, 155–162.

Berg, C. A., & Klaczynski, P. A. (1996). Practical intelligence and problem solving: Searching for perspectives. In F. Blanchard-Fields & T. M. Hess (Eds.), *Perspectives on cognitive change in adulthood and aging* (pp. 323–357). New York: McGraw-Hill.

Berg, K. M., Berg, W. K., & Graham, F. K. (1971). Infant heart rate response as a function of stimulus and state. *Psychophysiology*, 8, 30–44.

Berg, S. (1996). Aging, behavior, and terminal decline. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 323–337). San Diego: Academic Press.

Berg, W. K., Adkinson, C. D., & Strock, B. D. (1973). Duration and frequency of periods of alertness in neonates. *Developmental Psychology*, 9, 434.

Berg, W. K., & Berg, K. M. (1987). Psychophysiological development in infancy: State, startle, and attention. In J. D. Osofsky (Ed.), *Handbook of infant development* (2nd ed., pp. 238–317). New York: Wiley.

Berg-Cross, L. G. (1975). Intentionality, degree of damage, and moral judgments. *Child Development*, 46, 970–974.

Berkowitz, L., & Donnerstein, E. (1982). External validity is more than skin deep. *American Psychologist*, 37, 245–257.

Bertenthal, B. I., & Campos, J. J. (1990). A systems approach to the organizing effects of self-produced locomotion during infancy. In C. Rovee-Collier (Ed.), *Advances in infancy research* (Vol. 6, pp. 1–60). Norwood, NJ: Ablex.

Bertenthal, B. I., Campos, J. J., & Barrett, K. C. (1984). Self-produced locomotions: An organizer of emotional, cognitive, and social development in infancy. In R. Emde & R. Hartson (Eds.), *Continuities and discontinuities in development* (pp. 175–210). New York: Plenum.

Birren, J. E., & Morrison, D. F. (1961). Analysis of the WAIS subtests in relation to age and education. *Journal of Gerontology*, 16, 363–369.

Birren, J. E., Woods, A. M., & Williams, M. V. (1980). Behavioral slowing with age: Causes, organization, and consequences. In L. W. Poon (Ed.), *Aging in the 1980s: Psychological issues* (pp. 293–308). Washington, DC: American Psychological Association.

Bjorklund, D. F. (1987). How age changes in knowledge base contribute to the development of children's memory: An integrative review. *Developmental Review*, 7, 93–130.

Bjorklund, D. F., Coyle, T. R., & Gaultney, J. F. (1992). Developmental differences in the acquisition and maintenance of an organizational strategy: Evidence for the utilization deficiency hypothesis. *Journal of Experimental Child Psychology*, 54, 434–448.

Blair, R. C., & Higgins, J. J. (1980). A comparison of the power of Wilcoxon's rank sum statistic to that of student's t statistic under various nonnormal distributions. *Journal of Educational Statistics*, 5, 309–335.

Blasi, A. (1980). Bridging moral cognition and moral action. *Psychological Bulletin*, 88, 593–637.

Block, J. H. (1976). Issues, problems, and pitfalls in assessing sex differences: A critical review of *The psychology of sex differences*. *Merrill-Palmer Quarterly*, 22, 285–308.

Boldizar, J. P. (1991). Assessing sex typing and androgyny in children. *Developmental Psychology*, 27, 505–515.

Bornstein, M. H. (1978). Chromatic vision in infancy. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 12, pp. 117–182). New York: Academic Press.

Bornstein, M. H. (Ed.). (1995). *Handbook of parenting*. Hillsdale, NJ: Erlbaum.

Botwinick, J. (1984). *Aging and behavior* (3rd ed.). New York: Springer.

Botwinick, J., & Bitten, J. E. (1963). Cognitive processes: Mental abilities and psychomotor responses in healthy aged men. In J. E. Birren, R. N. Butler, S. W. Greenhouse, L. Sokoloff, & M. R. Yarrow (Eds.), *Human aging* (pp. 97–108). Washington, DC: U.S. Government Printing Office.

Botwinick, J., & Storandt, M. (1974). *Memory, related functions, and age*. Springfield, IL: Charles C. Thomas.

Botwinick, J., West, R. L., & Storandt, M. (1978). Predicting death from behavioral test performance. *Journal of Gerontology*, 33, 755–762.

Bouchard, T. J., Jr., Lykken, D. T., McGue, M., Segal, N. L., & Tellegen, A. (1990). Sources of human psychological differences: The Minnesota Study of Twins Reared Apart. *Science*, 250, 223–228.

Bower, T. G. R. (1966). The visual world of infants. *Scientific American*, 215, 90–92.

Bowlby, J. (1952). *Maternal care and mental health* (2nd ed.). Geneva: World Health Organization.

Braungart, J. M., Ploniin, R., DeFries, J. C., & Fulker, D. W. (1992). Genetic influence on tester-rated infant temperament as assessed by Bayley's Infant Behavior Record: Non-adoptive and adoptive siblings and twins. *Developmental Psychology*, 28, 40–47.

Brazelton, T. B., & Nugent, J. K. (1995). *Neonatal Behavioral Assessment Scale* (3rd ed.). Clinics in Developmental Medicine, No. 137. London: Mae Keith Press.

Brazelton, T. B., Nugent, J. K., & Lester, B. M. (1987). Neonatal Behavioral Assessment Scale. In J. D. Osofsky (Ed.), *Handbook of infant development* (2nd ed., pp. 780–817). New York: Wiley.

Bretherton, I., & Ainsworth, M. D. S. (1974). Responses of one-year-olds to a stranger in strange situation. In M. Lewis & L. A. Rosenblum (Eds.), *The origins of fear* (pp. 131–164). New York: Wiley.

Brewer, J., & Hunter, A. (1989). *Multimethod research: A synthesis of styles*. Newbury Park, CA: Sage.

Brody, G. H., Stoneman, Z., & Wheatley, P. (1984). Peer interaction in the presence and absence of observers. *Child Development*, 55, 1425–1428.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, 32, 513–531.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (1989). Ecological systems theory. In R. Vasta (Ed.), *Annals of child development: Vol. 6. Six theories of child development: Revised formulations and current issues* (pp. 187–249). Greenwich, CT: JAI Press.

Bronfenbrenner, U. (1993). The ecology of cognitive development: Research models and fugitive findings. In R. H. Wozniak & K. W. Fischer (Eds.), *Development in context* (pp. 3–44). Hillsdale, NJ: Erlbaum.

Brooks, P. H., & Kendall, E. D. (1982). Working with children. In R. Vasta (Ed.), *Strategies and techniques of child study* (pp. 325–343). New York: Academic Press.

Brown, A. L. (1982). Learning to learn how to read. In J. Langer & T. Smith-Burke (Eds.), *Reader meets author, bridging the gap: A psycholinguistic perspective* (pp. 26–54). Newark, NJ: Dell.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Bryant, B. K. (1982). An index of empathy for children and adolescents. *Child Development*, 53, 412–425.

Bucher, B., & Schneider, R. E. (1973). Acquisition and generalization of conservation by pre-schoolers, using operant training. *Journal of Experimental Child Psychology*, 16, 187–204.

Bullock, M. (1995, July/August). What's so special about a longitudinal study? *Psychological Science Agenda*, 8, 9–10.

Buros, O. K. (Ed.). (1978). *The eighth mental measurements yearbook*. Highland Park, NJ: The Gryphon Press.

Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, NJ: Erlbaum.

Cairns, R. B., Cairns, B. D., Neckerman, H. J., Ferguson, L. L., & Garipey, J.-L. (1989). Growth and aggression: 1. Childhood to early adolescence. *Developmental Psychology*, 25, 320–330.

Cairns, R. B., & Green, J. A. (1979). How to assess personality and social patterns: Observations or ratings? In R. B. Cairns (Ed.), *The analysis of social interactions: Methods, issues, and illustrations* (pp. 209–226). Hillsdale, NJ: Erlbaum.

Camp, C. J., West, R. L., & Poon, L. W. (1989). Recruitment practices for psychological research in gerontology. In M. P. Lawton & A. R. Herzog (Eds.), *Special research methods for gerontology* (pp. 163–189). Amityville, NY: Baywood.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Campos, J. J., Langer, A., & Krowitz, A. (1970). Cardiac responses on the visual cliff in prelocomotor infants. *Science*, 170, 196–197.

Canadian Psychological Association (1991). *Canadian Code of Ethics for Psychologists-Revised*. Old Chelsea, Quebec: Author.

Canestrari, R. E., Jr. (1968). Age changes in acquisition. In G. A. Talland (Ed.), *Human aging and behavior* (pp. 169–188). New York: Academic Press.

Carey, W. B., & McDevit, S. C. (1978). Revision of the Infant Temperament Questionnaire. *Pediatrics*, 61, 735–739.

Carter, D. B., & Levy, G. D. (1988). Cognitive aspects of early sex-role development: The influence of gender schemas on preschoolers' memories and preferences for sex-typed toys and activities. *Child Development*, 59, 782–792.

Casler, L. (1961). Maternal deprivation: A critical review of the literature. *Mono-graphs of the Society for Research in Child Development*, 26 (2, Serial No. 80).

Chandler, M., & Chapman, M. (Eds.). (1991). *Criteria for competence: Controversies in the conceptualization and assessment of children's abilities*. Hillsdale, NJ: Erlbaum.

Chandler, M., Greenspan, S., & Barenboim, C. (1973). Judgments of intentionality in response to videotaped and verbally presented moral dilemmas: The medium is the message. *Child Development*, 44, 315–320.

Chandler, M., & Hala, S. (1994). The role of personal involvement in the assessment of early false belief skills. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind* (pp. 403–425). Hillsdale, NJ: Erlbaum.

Cherry, K. E., & Park, D. C. (1993). Individual difference and contextual variables influence spatial memory in younger and older adults. *Psychology and aging*, 8, 517–526.

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, 9, 377–395.

Cliff, N. (1993). What is and isn't measurement. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 59–93). Hillsdale, NJ: Erlbaum.

Cohen, G. (1979). Language comprehension in old age. *Cognitive Psychology*, 11, 412–429.

Cohen, G. (1981). Inferential reasoning in old age. *Cognition*, 9, 59–72.

Cohen, G., Stanhope, N., & Conway, M. A. (1992). Age differences in the retention of knowledge by young and elderly students. *British Journal of Developmental Psychology*, 10, 153–164.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York: Academic Press.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 997–1003.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Erlbaum.

Cohen, S. H., & Reese, H. W. (Eds.). (1994). *Life-span developmental psychology: Methodological contributions*. Hillsdale, NJ: Erlbaum.

Colby, A., & Kohlberg, L. (1987). *The measurement of moral judgment: Vol. 1. Theoretical foundations and research validation*. New York: Cambridge University Press.

Colby, A., Kohlberg, L., Gibbs, J., Candee, D., Hewer, A., Power, C., & Speicher-Dubin, B. (1987). *The measurement of moral judgment: Vol. 2. Standard issue scoring manual*. New York: Cambridge University Press.

Colby, A., Kohlberg, L., Gibbs, J., & Lieberman, M. (1983). A longitudinal study of moral judgment. *Monographs of the Society for Research in Child Development*, 48 (1-2, Serial No. 200).

Cole, M., Frankel, F., & Sharp, D. (1971). Development of free recall learning in children. *Developmental Psychology*, 4, 109-123.

Collins, L. M., & Horn, J. L. (Eds.). (1991). *Best methods for the analysis of change: Recent advances, unanswered questions, future directions*. Washington, DC: American Psychological Association.

Colombo, J. (1993). *Infant cognition: Predicting later intellectual functioning*. Thousand Oaks, CA: Sage.

Committee for Ethical Conduct in Child Development Research (1990, Winter). Report from the Committee for Ethical Conduct in Child Development Research. *SRCD Newsletter*, pp. 5-7.

Condry, J., & Condry, S. (1976). Sex differences: A study of the eye of the beholder. *Child Development*, 47, 812-819.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-Experimentation*. Boston: Houghton Mifflin.

Cooke, R. A. (1982). The ethics and regulation of research involving children. In B. B. Wolman (Ed.), *Handbook of developmental psychology* (pp. 149-172). Englewood Cliffs, NJ: Prentice-Hall.

Coopersmith, S. (1967). *The antecedents of self-esteem*. San Francisco: W. H. Freeman.

Coopersmith, S. (1981). *Coopersmith Self-Esteem Inventory*. Palo Alto, CA: Consulting Psychologists Press, Inc.

Costa, P. T., Jr., & McCrae, R. R. (1989). Personality continuity and the changes of adult life. In M. Storandt & G. R. VandenBos (Eds.), *The adult years: Continuity and change* (pp. 45-77). Washington, DC: American Psychological Association.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Manual for Revised NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory (NEO-FFI)*. Odessa, FL: Psychological Assessment Resources, Inc.

Cox, B. D., Ornstein, P. A., Naus, M. J., Maxfield, D., & Zimler, J. (1989). Children's concurrent use of rehearsal and organizational strategies. *Developmental Psychology*, 25, 619-627.

Craik, F. I. M. (1994). Memory changes in normal aging. *Current Directions in Psychological Science*, 3, 155-158.

Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). New York: Harper & Row.

Crook, T. H. & Larrabee, G. J. (1990). A self-rating scale for evaluating memory in everyday life. *Psychology and Aging*, 5, 48-57.

Cunningham, W. R. & Haman, K. L. (1992). Intellectual functioning in relation to mental health. In J. E. Birren, R. B. Sloane, & G. D. Cohen (Eds.), *Handbook of mental health and aging* (2nd ed., pp. 339-354). San Diego: Academic Press.

Damon, W. (1980). Patterns of change in children's social reasoning: A two-year longitudinal study. *Child Development*, 51, 1010–1017.

Damon, W., & Hart, D. (1988). *Self-understanding in childhood and adolescence*. New York: Cambridge University Press.

Darwin, C. (1877). Biographical sketch of an infant. *Mind*, 2, 285–294.

Dawe, H. C. (1934). An analysis of two hundred quarrels of preschool children. *Child Development*, 5, 139–157.

Deaux, K. (1993). Commentary: Sorry, wrong number—a reply to Gentile's call. *Psychological Science*, 4, 125–126.

DeLoache, J. S., Cassidy, D. J., & Brown, A. L. (1985). Precursors of mnemonic strategies in very young children's memory. *Child Development*, 56, 125–137.

Deur, J. L., & Parke, R. D. (1970). Effects of inconsistent punishment on aggression in children. *Developmental Psychology*, 2, 403–411.

Diehl, M., Willis, S. L., & Schaie, K. W. (1995). Everyday problem solving in older adults: Observational assessment and cognitive correlates. *Psychology and Aging*, 10, 478–491.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 107, 542–575.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, 49, 71–75.

Division of Developmental Psychology of the American Psychological Association (1968, Spring). Ethical standards in psychological research. *Newsletter*, 1–3.

Dixon, R. A. (1989). Questionnaire research on metamemory and aging: Issues of structure and function. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 394–416). Cambridge, Eng.: Cambridge University Press.

Donaldson, G., & Horn, J. L. (1992). Age, cohort, and time developmental muddles: Easy in practice, hard in theory. *Experimental Aging Research*, 18, 213–222.

Dombusch, S. M., Ritter, P. L., Leiderman, P. H., Roberts, D. F., & Fraleigh, M. J. (1987). The relation of parenting style to adolescent school performance. *Child Development*, 58, 1244–1257.

Downs, A. C., & Langlois, J. H. (1988). Sex typing: Construct and measurement issues. *Sex Roles*, 18, 87–100.

Draper, N. R., & Smith, H. (1981). *Applied regression analyses* (2nd ed.). New York: Wiley.

Dufresne, A., & Kobasigawa, A. (1989). Children's spontaneous allocation of study time: Differential and sufficient aspects. *Journal of Experimental Child Psychology*, 47, 274–296.

Dunn, J. (1995). Children as psychologists: The later correlates of individual differences in understanding of emotions and other minds. *Cognition and Emotion*, 9, 187–201.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN: American Guidance Service.

Edelbrock, C., & Sugawara, A. I. (1978). Acquisition of sex-typed preferences in pre-school-aged children. *Developmental Psychology*, 14, 614–623.

Edelman, A., Kraemer, H., & Komer, A. (1982). Effects of compensatory movement stimulation on the sleep-wake behaviors of preterm infants. *Journal of the American Academy of Child Psychiatry*, 21, 555–559.

Eder, R. A. (1989). The emerging personologist: The structure and content of 2 1/2-, 5 1/2-, and 7 1/2b year-olds' concepts of themselves and others. *Child Development*, 60, 1218–1228.

Edwards, C. P. (1986). Cross-cultural research on Kohlberg's stages: The basis for consensus. In S. Modgil & C. Modgil (Eds.), *Lawrence Kohlberg: Consensus and controversy* (pp. 419–430). Philadelphia: Falmer.

Edwards, J., & Klemmack, D. (1973). Correlates of life satisfaction: A reexamination. *Journal of Gerontology*, 28, 497–502.

Eisenberg, N. (1982). The development of reasoning regarding prosocial behavior. In N. Eisenberg (Ed.), *The development of prosocial behavior* (pp. 219–249). New York: Academic Press.

Eisenberg, N., Fabes, R. A., Bustamante, D., Mathy, R. M., Miller, P. A., & Lindholm, E. (1988). Differentiation of vicariously induced emotional reactions in children. *Developmental Psychology*, 24, 237–246.

Eisenberg, N., Fabes, R. A., Miller, P. A., Fultz, J., Shell, R., Mathy, R. M., & Reno, R. R. (1989). Relation of sympathy and personal distress to prosocial behavior: A multi-method study. *Journal of Personality and Social Psychology*, 57, 55–66.

Elder, G. H. (1974). *Children of the Great Depression*. Chicago: University of Chicago Press.

Elder, G. H., & Caspi, A. (1990). Studying lives in a changing society: Sociological and personological explanations. In A. I. Rabin, R. A. Zucker, & S. Frank (Eds.), *Studying persons and lives* (pp. 201–247). New York: Springer.

Elias, M. F., Elias, J. W., & Elias, P. K. (1990). Biological and health influences on behavior. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 79–102). San Diego: Academic Press.

Elicker, J., Englund, M., & Sroufe, L. A. (1992). Predicting peer competence and peer relationships in childhood from early parent-child relationships. In R. D. Parke & G. W. Ladd (Eds.), *Family-peer relationships: Modes of linkage* (pp. 77–106). Hillsdale, NJ: Erlbaum.

Ellis, G. T., & Sekyra, F. (1972). The effect of aggressive cartoons on the behavior of first grade children. *Journal of Psychology*, 81, 37–43.

Emde, R. N., Plomin, R., Robinson, J., Corley, R., DeFries, J., Fulker, D. W., Reznick, J. S., Campos, J., Kagan, J., & Zahn-Waxler, C. (1992). Temperament, emotion, and cognition at fourteen months: The MacArthur Longitudinal Twin Study. *Child Development*, 63, 1437–1455.

Emmerich, W., & Goldman, K. S. (1972). Boy-girl identity task. In V. Shipman (Ed.), *Disadvantaged children and their first school experiences (ETS PR-72-20)*. Princeton, NJ: Educational Testing Service.

Engen, T., Lipsitt, L. P., & Kaye, H. (1963). Olfactory responses and adaptation in the human neonate. *Journal of Comparative and Physiological Psychology*, 56, 73–77.

Entwisle, D. R., & Astone, N. M. (1994). Some practical guidelines for measuring youth's race/ethnicity and socioeconomic status. *Child Development*, 65, 1521–1540.

Erber, J. T. (1981). Remote memory and age: A review. *Experimental Aging Research*, 7, 189–200. Fagan, J. F., III, & Detterman, D. H. (1992). The Fagan Test of Infant Intelligence: A technical summary. *Journal of Applied Developmental Psychology*, 10, 173–193.

Fagan, J. F., III, & Shepherd, P. A. (1986). *The Fagan Test of Infant Intelligence: Training manual*. Cleveland: Infantest Corporation.

Fagot, B. I. (1977). Consequences of moderate crossgender behavior in preschool children. *Child Development*, 48, 902–907.

Fagot, B. I. (1985). Changes in thinking about early sex role development. *Developmental Review*, 5, 83–98.

Fagot, B. I., & Leinbach, M. D. (1993). Gender-role development in young children: From discrimination to labeling. *Developmental Review*, 13, 205–224.

Fantz, R. L. (1961). The origin of form perception. *Scientific American*, 204, 66–72.

Ferguson, R. P., & Bray, N. W. (1976). Component processes of an overt rehearsal strategy in young children. *Journal of Experimental Child Psychology*, 21, 490–506.

Feshbach, N. D., & Roe, K. (1968). Empathy in six- and seven-year-olds. *Child Development*, 39, 133–145.

Feshbach, S., & Singer, R. D. (1971). *Television and aggression: An experimental field study*. San Francisco: Jossey-Bass.

Field, D. (1987). A review of preschool conservation training: An analysis of analyses. *Developmental Review*, 7, 210–251.

Field, J. (1977). Coordination of vision and prehension in young infants. *Child Development*, 48, 97–103.

Field, T. (1982). Infancy. In R. Vasta (Ed.), *Strategies and techniques of child study* (pp. 13–48). New York: Academic Press.

Fillenbaum, G. G. (1978). Reliability and validity of the OARS Multidimensional Functional Assessment Questionnaire. In Duke University Center for the Study of Aging (Eds.), *Multidimensional functional assessment: The OARS methodology* (2nd ed., pp. 20–28). Durham, NC: Duke University.

Fillenbaum, G. G. (1985). Screening the elderly: A brief instrumental activities of daily living measure. *Journal of the American Geriatrics Society*, 33, 698–706.

Fisher, C. B., & Lerner, R. M. (Eds.). (1994a). *Applied developmental psychology*. New York: McGraw-Hill.

Fisher, C. B., & Lerner, R. M. (1994b). Foundations of applied developmental psychology. In C. B. Fisher & R. M. Lerner (Eds.), *Applied developmental psychology* (pp. 3–20). New York: McGraw-Hill.

Fisher, C. B., & Tryon, W. W. (Eds.). (1990). *Ethics in applied developmental psychology: Emerging issues in an emerging field*. Norwood, NJ: Ablex.

Flavell, J. H. (1963). *The developmental psychology of Jean Piaget*. Princeton, NJ: Van Nostrand.

Flavell, J. H. (1985). *Cognitive development* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Flavell, J. H. (1992). Perspectives on perspective taking. In H. Beilin & P. Pufall (Eds.), *Piaget's theory: Prospects and possibilities* (pp. 107–139). Hillsdale, NJ: Erlbaum.

Flavell, J. H., Beach, D. H., & Chinsky, J. M. (1966). Spontaneous verbal rehearsal in memory tasks as a function of age. *Child Development*, 37, 283–299.

Flavell, J. H., Botkin, P. T., Fry, C. L., Jr., Wright, J. W., & Jarvis, P. E. (1968). *The development of role-taking and communication skills in children*. New York: Wiley.

Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance-reality distinction. *Cognitive Psychology*, 15, 95–120.

Flavell, J. H., Friedrichs, A. G., & Hoyt, J. (1970). Developmental changes in memorization processes. *Cognitive Psychology*, 1, 324–340.

Flavell, J. H., Miller, P. H., & Miller, S. A. (1993). *Cognitive development* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Fox, N. A., & Fitzgerald, H. E. (1990). Autonomic function in infancy. *Merrill-Palmer Quarterly*, 36, 27–52.

Freedman, J. L. (1984). Effect of television violence on aggressiveness. *Psychological Bulletin*, 96, 227–246.

Gallup, G. G., Jr. (1970). Chimpanzees: Self-recognition. *Science*, 167, 86–87.

Gallup, G. G., Jr. (1994). Self-recognition: Research strategies and experimental design. In S. Parker, R. W. Mitchell, & M. L. Boccia (Eds.), *Self-awareness in animals and humans* (pp. 35–50). New York: Cambridge University Press.

Garbarino, J. (1982). Sociocultural risk: Dangers to competence. In C. B. Kopp & J. B. Krakow (Eds.), *The child: Development in a social context* (pp. 630–685). Reading, MA: Addison-Wesley.

Gelman, R. (1969). Conservation acquisition: A problem of learning to attend to relevant attributes. *Journal of Experimental Child Psychology*, 7, 167–187.

Gelman, R. (1972). Logical capacity of very young children: Number invariance rules. *Child Development*, 43, 75–90.

Gelman, R. (1991). Epigenetic foundations of knowledge structures: Initial and transcendent constructions. In S. Carey & R. Gehlan (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 293–322). Hillsdale, NJ: Erlbaum.

Gentile, D. A. (1993). Just what are sex and gender, anyway? A call for a new terminological standard. *Psychological Science*, 4, 120–122.

Gibbons, J. D. (1993). *Nonparametric statistics: An introduction*. Newbury Park, CA: Sage.

Gibbs, E. D., & Teti, D. M. (Eds.). (1990). *Interdisciplinary assessment of infants*. Baltimore: Paul H. Brookes Publishing Co.

Gibbs, J. C., Basinger, K. S., & Fuller, D. (1992). *Moral maturity: Measuring the development of sociomoral reflection*. Hillsdale, NJ: Erlbaum.

Gibbs, J. C., Widaman, K. F., & Colby, A. (1982). Construction and validation of a simplified, group-administrable equivalent to the moral judgment interview. *Child Development*, 53, 895–910.

Gilewski, M. J., & Zelinski, E. M. (1986). Questionnaire assessment of memory complaints. In L. W. Poon, T. H. Crook, K. L. Davis, C. Eisdorfer, B. J. Gurland, A. W. Kasniak, & L. W. Thompson (Eds.), *Handbook for clinical memory assessment of older adults* (pp. 93–107). Washington, DC: American Psychological Association.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.

Ginsburg, H., & Oppen, S. (1988). *Piaget's theory of intellectual development: An introduction* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Goldsmith, H. H., & Alansky, J. A. (1987). Maternal and infant temperamental predictors of attachment: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 55, 805–816.

Goldsmith, H. H., Buss, A. H., Plomin, R., Rothbart, M. K., Thomas, A., Chess, S., Hinde, R. A., & McCall, R. B. (1987). Roundtable: What is temperament: Four approaches. *Child Development*, 58, 505–529.

Goldsmith, H. H., & Rothbart, M. K. (1991). Contemporary instruments for assessing early temperament by questionnaire and in the laboratory. In J. Strelan & A. Angleitner (Eds.), *Explorations in temperament: International perspectives on theory and measurement* (pp. 249–272). New York: Plenum Press.

Goldsmith, H. H., & Rothbart, M. K. (1992). *The Laboratory Temperament Assessment Battery*. Eugene, OR: Personality Development Laboratory.

Golombok, S., & Fivush, R. (1994). *Gender development*. Cambridge, Eng.: Cambridge University Press.

Golombok, S., & Rust, J. (1993a). The measurement of gender role behavior in preschool children: A research note. *Journal of Child Psychology and Psychiatry*, 34, 805–811.

Golombok, S., & Rust, J. (1993b). The Pre-School Activities Inventory: A standardized assessment of gender role in children. *Psychological Assessment*, 5, 131–136.

Gonda, J. (1980). Relationship between formal education and cognitive functioning: A historical perspective. *Educational Gerontology*, 5, 283–291.

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, 59, 26–37.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Gottlieb, G., & Krasnegor, N. A. (Eds.). (1984). *Measurement of audition and vision in the first year of postnatal life: A methodological overview*. Norwood, NJ: Ablex.

Graham, S. (1992). «Most of the subjects were white and middle class»: Trends in published research on African Americans in selected APA journals, 1970–1989. *American Psychologist*, 47, 629–639.

Gray, S. W., Ramsey, B. K., & Klaus, R. A. (1982). *From 3 to 20: The Early Training Project*. Baltimore: University Park Press.

Green, R. F. (1969). Age-intelligence relationship between ages sixteen and sixty-four: A rising trend. *Developmental Psychology*, 1, 618–627.

Gruen, G. E. (1965). Experiences affecting the development of number conservation. *Child Development*, 36, 963–979.

Grusec, J. E. (1982). Prosocial behavior and self-control. In R. Vasta (Ed.), *Strategies and techniques of child study* (pp. 245–272). New York: Academic Press.

Grusec, J. E., & Kruczynski, L. (1980). Direction of effect in socialization: A comparison of the parent's versus the child's behavior as determinants of disciplinary techniques. *Developmental Psychology*, 16, 1–9.

Grusec, J. E., & Lytton, H. (1988). *Social development*. New York: Springer-Verlag.

Hagen, J. W., & Conley, A. C. (1994, Spring). Ethnicity and race of children studied in Child Development, 1980–1993. *SRCD Newsletter*, pp. 6–7.

Hainline, L., & Lemerise, E. (1982). Infants' scanning of geometric forms varying in size. *Journal of Experimental Child Psychology*, 33, 235–256.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1992). *Multivariate data analysis*. New York: MacMillan.

Haith, M. M. (1990). Progress in the understanding of sensory and perceptual processes in early infancy. *Merrill-Palmer Quarterly*, 36, 1–26.

Hall, J. A., & Halberstadt, A. G. (1980). Masculinity and femininity in children: Development of the Children's Personal Attributes Questionnaire. *Developmental Psychology*, 16, 270–280.

Hall, R. V., Fox, R., Willard, D., Goldsmith, L., Emerson, M., Owen, M., Davis, T., & Porcia, E. (1971). The teacher as observer and experimenter in the modification of disputing and talking-out behaviors. *Journal of Applied Behavior Analysis*, 4, 141–149.

Halverson, C. F., Jr., Kohnstamm, G. A., & Martin, R. P. (Eds.). (1994). *The developing structure of temperament and personality from infancy to adulthood*. Hillsdale, NJ: Erlbaum.

Harlow, H. F. (1958). The nature of love. *American Psychologist*, 13, 673–685.

Harlow, H. F. (1962). Fundamental principles for preparing psychology journal articles. *Journal of Comparative and Physiological Psychology*, 55, 893–896.

Harter, S. (1985). *The Self-Perception Profile for Children*. Denver, CO: University of Denver.

Harter, S. (1988). Developmental processes in the construction of the self. In T. D. Yawkey & J. E. Johnson (Eds.), *Integrative processes in socialization: Early to middle childhood* (pp. 45–78). Hillsdale, NJ: Erlbaum.

Harter, S., & Pike, R. (1984). The Pictorial Scale of Perceived Competence and Social Acceptance for Young Children. *Child Development*, 55, 1969–1982.

Hartmann, D. P. (1982). Assessing the dependability of observational data. In D. P. Hartmann (Ed.), *New directions for methodology of social and behavioral science: No. 14. Using observers to study behavior* (pp. 51–65). San Francisco: Jossey-Bass.

Hartmann, D. P. (1992). Design, measurement, and analysis: Technical issues in developmental research. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental psychology: An advanced textbook* (3rd ed., pp. 59–151). Hillsdale, NJ: Erlbaum.

Hartmann, D. P., & Wood, D. W. (1990). Observational methods. In A. S. Bellack, M. Hersen, & A. E. Kazdin (Eds.), *International handbook of behavior modification and therapy* (2nd ed., pp. 107–138). New York: Plenum.

Hartshorne, H., & May, M. A. (1928). *Studies in the nature of character: Vol. 1. Studies in deceit*. New York: Macmillan.

Hays, W. L. (1981). *Statistics* (3rd ed.). New York: Holt, Rinehart & Winston.

Helson, R. (1993). Comparing longitudinal studies of adult development: Toward a paradigm of tension between stability and change. In D. C. Funder, R. D. Parke, C. Tomlinson-Keasey, & K. Widarnan (Eds.), *Studying lives through time* (pp. 93–119). Washington, DC: American Psychological Association.

Hertzog, C. (1996). Research designs in studies of aging and cognition. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 24–37). San Diego: Academic Press.

Hertzog, C., & Dixon, R. A. (1996). Methodological issues in research on cognition and aging. In F. Blanchard-Fields & T. M. Hess (Eds.), *Perspectives on cognitive change in adulthood and aging* (pp. 66–121). New York: McGraw-Hill.

Hertzog, C., Schaie, K. W., & Gribbin, K. (1978). Cardiovascular disease and changes in intellectual functioning from middle to old age. *Journal of Gerontology*, 33, 872–883.

Hess, T. M., & Pullen, S. M. (1996). Memory in context. In F. Blanchard-Fields & T. M. Hess (Eds.), *Perspectives on cognitive change in adulthood and aging* (pp. 387–427). New York: McGraw-Hill.

Hetherington, E. M., & Frankie, G. (1967). Effects of parental dominance, warmth, and conflict on imitation in children. *Journal of Personality and Social Psychology*, 6, 119–125.

Hicks, D. J. (1968). Effects of co-observer's sanctions and adult presence on imitative aggression. *Child Development*, 39, 303–309.

Hoffman, M. L. (1975). Sex differences in moral internalization and values. *Journal of Personality and Social Psychology*, 32, 720–729.

Hoffman, M. L., & Saltzstein, H. D. (1967). Parent discipline and the child's moral development. *Journal of Personality and Social Psychology*, 5, 45–57.

Hogrefe, G. J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development*, 57, 567–582.

Holden, G. W., & Edwards, L. A. (1989). Parental attitudes toward child rearing: Instruments, issues, and implications. *Psychological Bulletin*, 106, 29–58.

Hollander, M., & Wolfe, D. A. (1973). *Nonparametric statistical methods*. New York: Wiley.

Hollenbeck, A. R. (1978). Problems of reliability in observational research. In G. P. Sackett (Ed.), *Observing behavior: Vol. 2. Data collection and analysis methods* (pp. 79–98). Baltimore: University Park Press.

Holmes, H. A., Black, C., & Miller, S. A. (1996). A cross-task comparison of false belief understanding in a Head Start population. *Journal of Experimental Child Psychology*, 63, 263–285.

Horka, S., & Farrow, B. (1970). A methodological note on intersubject communication as a contaminating factor in psychological experiments. *Journal of Experimental Child Psychology*, 10, 363–366.

Hom, J. L. (1982). The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood. In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 237–278). New York: Plenum.

Howes, J. L., & Katz, A. N. (1988). Assessing remote memory with an improved public events questionnaire. *Psychology and Aging*, 3, 142–150.

Hubert, N. C., Wachs, T. D., Peters-Martin, P., & Gandour, M. J. (1982). The study of early temperament: Measurement and conceptual issues. *Child Development*, 53, 571–600.

Hudson, J. A., & Nelson, K. (1983). Effects of script structure on children's story recall. *Developmental Psychology*, 19, 625–635.

Huesmann, L. R., Lagerspetz, K., & Eton, L. D. (1984). Intervening variables in the TV violence-aggression relation: Evidence from two countries. *Developmental Psychology*, 20, 746–775.

Huston, A. C., McLoyd, V. C., & Coll, C. G. (Eds.). (1994). Children and poverty [Special issue]. *Child Development*, 65 (2).

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child*. London: Routledge & Kegan Paul.

Izard, C. E. (1979). *The Maximally Discriminative Facial Movement Coding System* (Max.) Newark: University of Delaware, Instructional Resources Center.

Jeffrey, D. B., Hartmann, D. P., & Gelfand, D. M. (1972). A comparison of the effects of contingent reinforcement, nurturance, and nonreinforcement on initiative learning. *Child Development*, 43, 1053–1059.

Jensen, A. R. (1981). *Straight talk about mental tests*. New York: Free Press.

Kagan, J. (1964). American longitudinal research on psychological development. *Child Development*, 35, 1–32.

Kagan, J., Reznick, J. S., & Gibbons, J. (1989). Inhibited and uninhibited types of children. *Child Development*, 60, 838–845.

Kail, R. V., & Cavanaugh, J. C. (1996). *Human development*. Pacific Grove, CA: Brooks/Cole.

- Kamin, L. J. (1974). *The science and politics of IQ*. Hillsdale, NJ: Erlbaum.
- Kausler, D. H. (1994). *Learning and memory in normal aging*. New York: Academic Press.
- Kazdin, A. E. (1992). *Research design in clinical psychology* (2nd ed.). New York: Macmillan.
- Keeney, T., Cannizzo, S. R., & Flavell, J. H. (1967). Spontaneous and induced verbal rehearsal in a recall task. *Child Development*, 38, 953–966.
- Keith, L. R., & Bracken, B. A. (1996). Self-concept instrumentation: A historical and evaluative review. In B. A. Bracken (Ed.), *Handbook of self-concept: Developmental, social, and clinical considerations* (p. 91–170). New York: Wiley.
- Kent, R. N., O'Leary, K. D., Diamant, C., & Dietz, A. (1974). Expectation biases in observational evaluation of therapeutic change. *Journal of Consulting and Clinical Psychology*, 42, 774–780.
- Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Kerlinger, F. N. (1986). *Foundations of behavioral research* (3rd ed.). New York: Holt, Rinehart & Winston.
- Kirasic, K. C. (1991). Spatial cognition and behavior in young and elderly adults: Implications for learning new environments. *Psychology and Aging*, 6, 10–18.
- Klein, R. L. (1972). Age, sex, and task difficulty as predictors of social conformity. *Journal of Gerontology*, 27, 229–236.
- Kofsky, E. (1966). A scalogram study of classificatory development. *Child Development*, 37, 191–204.
- Kohlberg, L. (1966). A cognitive-developmental analysis of children's sex-role concepts and attitudes. In E. Maccoby (Ed.), *The development of sex differences* (pp. 82–173). Stanford, CA: Stanford University Press.
- Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347–480). Chicago: Rand McNally.
- Kopp, C. B. (1994). Infant assessment. In C. B. Fisher & R. M. Lerner (Eds.), *Applied developmental psychology* (pp. 265–293). New York: McGraw-Hill.
- Korner, A. F., & Thoman, E. B. (1970). Visual alertness in neonates as evoked by maternal care. *Journal of Experimental Child Psychology*, 10, 67–78.
- Korner, A. F., & Thoman, E. B. (1972). The relative efficacy of contact and vestibular-proprioceptive stimulation in soothing neonates. *Child Development*, 43, 443–453.
- Kuczynski, L. (1983). Reasoning, prohibitions, and motivations for compliance. *Developmental Psychology*, 19, 126–134.
- Kurtines, W., & Greif, E. B. (1974). The development of moral thought: Review and evaluation of Kohlberg's approach. *Psychological Bulletin*, 81, 453–470.

Labouvie, E. W. (1980). Identity vs. equivalence of psychological measures and constructs. In L. W. Poon (Ed.), *Aging in the 1980s: Psychological issues* (pp. 493–502). Washington, DC: American Psychological Association.

LaGreca, A. M. (Ed.). (1990). *Through the eyes of the child: Obtaining self-reports from children and adolescents*. Needham Heights, MA: Allyn & Bacon.

Lamb, M. E. (1976). Twelve-month-olds and their parents: Interaction in a laboratory playroom. *Developmental Psychology*, 12, 237–244.

Lamb, M. E., Easterbrooks, M. A., & Holden, G. W. (1980). Reinforcement and punishment among preschoolers: Characteristics, effects, and correlates. *Child Development*, 51, 1230–1236.

Lamb, M. E., & Roopnarine, J. L. (1979). Peer influences on sex-role development in preschoolers. *Child Development*, 50, 1219–1222.

Lamborn, S. D., Mounts, N. S., Steinberg, L., & Dombusch, S. M. (1991). Patterns of competence and adjustment among adolescents from authoritative, authoritarian, indulgent, and neglectful families. *Child Development*, 62, 1049–1065.

Langford, P. E. (1995). *Approaches to the development of moral reasoning*. Hove, UK: Erlbaum.

LaRue, A., Bank, L., Jarvik, L., & Hetland, M. (1979). Health in old age: How do physicians' ratings and self-ratings compare? *Journal of Gerontology*, 34, 687–691.

Laurence, M. W. (1967). Memory loss with age: A test of two strategies for its retardation. *Psychonomic Science*, 9, 209–210.

Lawton, M. P., Whelihan, W. M., & Belsky, J. K. (1980) Personality tests and their uses with older adults. In J. E. Birren & R. B. Sloane (Eds.), *Handbook of mental health and aging* (pp. 537–553). Englewood Cliffs, NJ: Prentice-Hall.

Lefkowitz, M. M., Eton, L. D., Walder, L. O., & Huesmann, L. R. (1972). Television violence and child aggression: A followup study. In G. A. Comstock & E. A. Rubinstein (Eds.), *Television and social behavior: Vol. 3. Television and adolescent aggressiveness* (pp. 35–135). Washington, DC: U.S. Government Printing Office.

Leinbach, M. D., & Fagot, B. I. (1986). Acquisition of gender labeling: A test for toddlers. *Sex Roles*, 15, 655–666.

Levin, J. R. (1985). Some methodological and statistical «bugs» in research on children's learning. In M. Pressley & C. J. Brainerd (Eds.), *Cognitive learning and memory in children: Progress in cognitive development research* (pp. 205–233). New York: Springer-Verlag.

Lewis, C., & Mitchell, P. (Eds.). (1994). *Children's early understanding of mind*. Hove, UK: Erlbaum.

Lewis, M., & Johnson, N. (1971). What's thrown out with the bath water: A baby? *Child Development*, 42, 1053–1055.

Liben, L. S., & Signorella, M. L. (1993). Gender-schematic processing in children: The role of initial interpretations of stimuli. *Developmental Psychology*, 29, 141–149.

Liebert, R. M., & Baron, R. A. (1972). Some immediate effects of televised violence on children's behavior. *Developmental Psychology*, 6, 469–475.

Light, P. H., Buckingham, N., & Robbins, A. H. (1979). The conservation task as an interactional setting. *British Journal of Educational Psychology*, 49, 304–310.

Lipsitt, L. P. (1992). Discussion: The Bayley Scales of Infant Development: Issues of prediction and outcome revisited. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 7, pp. 239–245). Norwood, NJ: Ablex.

Livesley, W. J., & Bromley, D. B. (1973). *Person perception in childhood and adolescence*. London: Wiley.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

Main, M., & Solomon, J. (1991). Procedures for identifying infants as disorganized/disoriented during the Ainsworth Strange Situation. In M. Greenberg, D. Cicchetti, & E. M. Cummings (Eds.), *Attachment in the preschool years: Theory, research, and intervention* (pp. 121–160). Chicago: University of Chicago Press.

Mandler, J. M. (1983). Representation. In P. H. Mussen (Series Ed.) & J. H. Flavell & E. M. Markham (Vol. Eds.), *Handbook of child psychology: Vol. 3. Cognitive development* (4th ed., pp. 420–494). New York: Wiley.

Mann, J., Have, J., Plunkett, J., & Meisels, S. (1991). Time sampling: A methodological critique. *Child Development*, 62, 227–241.

Marascuilo, L. A., & McSweeney, M. (1977). *Nonparametric and distribution free methods for the social sciences*. Monterey, CA: Brooks/Cole.

Marcus, D. E., & Overton, W. F. (1978). The development of cognitive gender constancy and sex role preferences. *Child Development*, 49, 434–444.

Marsh, H. W. (1988). *Self-Description Questionnaire, I*. San Antonio, TX: The Psychological Corporation.

Marsiske, M., & Willis, S. L. (1995). Dimensionality of everyday problem solving in older adults. *Psychology and Aging*, 10, 269–283.

Martin, C. L. (1993). New directions for investigating children's gender knowledge. *Developmental Review*, 13, 184–204.

Martino-Salzman, D., Blasch, B. B., Morris, R. D., & McNeal, L. W. (1991). Travel behavior of nursing home residents perceived as wanderers and nonwanderers. *Gerontologist*, 31, 666–672.

May, C. P., Hasher, L., & Stolzhus, E. R. (1993). Optimal time of day and the magnitude of age differences in memory. *Psychological Science*, 4, 326–330.

McCall, R. B. (1977). Challenges to a science of developmental psychology. *Child Development*, 48, 333–344.

McCall, R. B., & Mash, C. W. (1994). Infant cognition and its relation to mature intelligence. In R. Vasta (Ed.), *Annals of child development* (Vol. 10, pp. 27–56). London: Kingsley.

McCall, R. B., Parke, R. D., & Kavanaugh, R. D. (1977). Imitation of live and televised models by children one to three years of age. *Monographs of the Society for Research in Child Development*, 42 (5, Serial No. 173).

McCartney, K., & Nelson, K. (1981). Children's use of scripts in story recall. *Discourse Processes*, 4, 59–70.

McClowry, S. G. (1995). The development of the *School-Age Temperament Inventory*, *MerrillPalmer Quarterly*, 41, 272–285.

McCormack, P. D. (1981). Temporal coding by young and elderly adults: A test of the Hasher-Zacks model. *Developmental Psychology*, 17, 509–515.

McCrae, R. R., & Costa, P. T., Jr. (1990). *Personality in adulthood*. New York: Guilford.

McCrae, R. R., & Costa, P. T., Jr. (1994). The stability of personality: Observations and evaluations. *Current Directions in Psychological Science*, 3, 173–175.

McDonough, L., & Mandler, J. M. (1989, April). *Immediate and deferred imitation with 11-month-olds: A comparison between familiar and novel actions*. Paper presented at the meeting of the Society for Research in Child Development, Kansas City, Mo.

McGarrigle, J., & Donaldson, M. (1974). Conservation accidents. *Cognition*, 3, 341–350.

McLeod, J. M., Atkin, C. K., & Chaffee, S. H. (1972). Adolescents, parents, and television use: Adolescent self-report measures from Maryland and Wisconsin samples. In G. A. Comstock & E. A. Rubinstein (Eds.), *Television and social behavior: Vol. 3. Television and adolescent aggressiveness* (pp. 173–238). Washington, DC: U.S. Government Printing Office.

McLoyd, V. C., & Spencer, M. B. (Eds.). (1990). Minority children [special issue]. *Child Development*, 61 (2).

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of personality*, 40, 525–543.

Meltzoff, A. N. (1988a). Infant imitation after a 1-week delay: Long-term memory for novel and multiple stimuli. *Developmental Psychology*, 24, 470–476.

Meltzoff, A. N. (1988b). Infant imitation and memory: Nine-month-olds in immediate and deferred tests. *Child Development*, 59, 217–225.

Messick, S. (1983). Assessment of children. In P. H. Mussen (Series Ed.) & W. Kessen (Vol. Ed.), *Handbook of child psychology: Vol. 1. History, theory, and methods* (4th ed., pp. 477–526). New York: Wiley.

Michell, J. (1986). Measurement scales and statistics: A clash of paradigms. *Psychological Bulletin*, 100, 398–407.

Miller, P. H. (1990). The development of strategies of selective attention. In D. F. Bjorklund (Ed.), *Children's strategies: Contemporary views of cognitive development* (pp. 157–184). Hillsdale, NJ: Erlbaum.

Miller, S. A. (1976a). Nonverbal assessment of conservation of number. *Child Development*, 47, 722–728.

Miller, S. A. (1976b). Nonverbal assessment of Piagetian concepts. *Psychological Bulletin*, 83, 405–430.

Miller, S. A. (1977). A disconfirmation of the quantitative identity-quantitative equivalence sequence. *Journal of Experimental Child Psychology*, 24, 180–189.

Miller, S. A. (1982). Cognitive development: A Piagetian perspective. In R. Vasta (Ed.), *Strategies and techniques of child study* (pp. 161–207). New York: Academic Press.

Mitchell, S. K. (1979). Interobserver agreement, reliability, and generalizability of data collected in observational studies. *Psychological Bulletin*, 86, 376–390.

Moely, B. E., Olson, F. A., Hawles, T. G., & Flavell, J. H. (1969). Production deficiency in young children's clustered recall. *Developmental Psychology*, 1, 26–34.

Moen, P., Elder, G. H., & Luscher, K. (Eds.). (1995). *Examining lives in context: Perspectives on the ecology of human development*. Washington, DC: American Psychological Association.

Moran, J. D., III, & McCullers, J. C. (1984). The effects of recency and story content on children's moral judgments. *Journal of Experimental Child Psychology*, 38, 447–455.

Morrison, D. F. (1990). *Multivariate statistical methods* (3rd ed.). New York: McGraw-Hill.

Moses, L. J., & Flavell, J. H. (1990). Inferring false beliefs from actions and reactions. *Child Development*, 61, 929–945.

Mussen, P. H., & Eisenberg-Berg, N. (1977). *Roots of caring, sharing, and helping: The development of prosocial behavior in children*. San Francisco: W. H. Freeman.

Myers, D. G., & Diener, E. (1995). Who is happy? *Psychological Science*, 6, 10–19.

Neale, J. M., & Liebert, R. M. (1986). *Science and behavior: An introduction to methods of research* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Nelson, E. A., & Daneffer, D. (1992). Aged heterogeneity: Fact or fiction: The fate of diversity in gerontological research. *Gerontologist*, 32, 17–23.

Nelson-LeGall, S. (1985). Motive-outcome matching and outcome foreseeability: Effects on attribution of intentionality and moral judgments. *Developmental Psychology*, 21, 332–337.

Nesselroade, J. R., & Baltes, P. B. (1974). Adolescent personality development and historical change: 1970–1972. *Monographs of the Society for Research in Child Development*, 39, (1, Serial No. 154).

Nesselroade, J. R., & Cattell, R. B. (Eds.). (1988). *Handbook of multivariate experimental psychology* (2nd ed.). London: Plenum Press.

Neugarten, B. L., Havighurst, R. J., & Tobin, S. S. (1961). The measurement of life satisfaction. *Journal of Gerontology*, 16, 134–143.

Nicholls, R. C. (1978). Heredity and environment: Major findings from twin studies of ability, personality, and interests. *Homo*, 29, 158–173.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Nunnally, J. C. (1982). The study of human change: Measurement, research strategies, and methods of analysis. In B. B. Wolman (Ed.), *Handbook of developmental psychology* (pp. 133–148). Englewood Cliffs, NJ: Prentice-Hall.

Odom, S. L., & Ogawa, I. (1992). Direct observation of young children's social interaction with peers: A review of methodology. *Behavioral Assessment*, 14, 407-441.

Ohta, R. J. (1981). Spatial problem-solving: The response selection tendencies of young and elderly adults. *Experimental Aging Research*, 1, 81-84.

Okun, M. A., & Elias, C. S. (1977). Cautiousness in adulthood as a function of age and payoff structure. *Journal of Gerontology*, 32, 451-455.

Olson, G. M., & Sherman, T. (1983). Attention, learning, and memory in infants. In P. H. Mussen (Series Ed.) & M. M. Haith & J. J. Campos (Vol. Eds.), *Handbook of child psychology: Vol. 2. Infancy and developmental psychobiology* (4th ed., pp. 1001-1080). New York: Wiley.

O'Neill, D. K., Astington, J. W., & Flavell, J. H. (1992). Young children's understanding of the role that sensory experiences play in knowledge acquisition. *Child Development*, 63, 474-490.

O'Neill, D. K., & Gopnik, A. (1991). Young children's ability to identify the sources of their beliefs. *Developmental Psychology*, 27, 390-397.

Oppenheim, D., & Waters, H. S. (1995). Narrative processes and attachment representations: Issues of development and assessment. In E. Waters, B. E. Vaughn, G. Posada, & K. Kondo-Ikemura (Eds.), *Caregiving, cultural, and cognitive perspectives on secure-base behavior and working models* (pp. 197-215). *Monographs of the Society for Research in Child Development*, 60 (2-3, Serial No. 244).

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17, 776-783.

Owens, W. A., Jr. (1966). Age and mental ability: A second follow-up. *Journal of Educational Psychology*, 57, 311-325.

Paris, S. G. (1975). Integration and inference in children's comprehension and memory. In F. Restle, R. Shiffrin, I. Castellan, H. Lindman, & D. Pisoni (Eds.), *Cognitive theory* (Vol. 1, pp. 223-246). Hillsdale, NJ: Erlbaum.

Paris, S. G., & Carter, A. Y. (1973). Semantic and constructive aspects of sentence memory in children. *Developmental Psychology*, 9, 109-113.

Paris, S. G., & Upton, L. R. (1976). Children's memory for inferential relationships in prose. *Child Development*, 47, 660-668.

Parke, R. D. (1967). Nurturance, nurturance withdrawal, and resistance to deviation. *Child Development*, 38, 1101-1110.

Parke, R. D. (1979). Interactional designs. In R. B. Cairns (Ed.), *The analysis of social interactions: Methods, issues, and illustrations* (pp. 15-35). Hillsdale, NJ: Erlbaum.

Parke, R. D. (1981). Some effects of punishment on children's behavior-revisited. In E. M. Hetherington & R. D. Parke (Eds.), *Contemporary readings in child psychology* (2nd ed., pp. 176-188). New York: McGraw-Hill.

Parke, R. D., Berkowitz, L., Leyens, J. P., West, S. G., & Sebastian, R. J. (1977). Some effects of violent and nonviolent movies on the behavior of juvenile delinquents. In

L. Berkowitz, (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 135–172). New York: Academic Press.

Parke, R. D., & Murray, S. (1971). *Re-instatement: A technique for increasing stability of inhibition in children*. Unpublished manuscript, University of Wisconsin.

Parrott, L., HI (1994). *How to write psychology papers*. New York: Harper Collins.

Patterson, C. J., & Carter, D. B. (1979). Attentional determinants of children's self-control in waiting and working situations. *Child Development*, 50, 272–275.

Patterson, G. R. (1979). A performance theory for coercive family interaction. In R. B. Cairns (Ed.), *The analysis of social interactions: Methods, issues, and illustrations* (pp. 119–162). Hillsdale, NJ: Erlbaum.

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia Press.

Patterson, G. R., & Bank, L. (1989). Some amplifying mechanisms for pathologic processes in families. In M. R. Gunnar & E. Thelen (Eds.), *Minnesota symposia on child psychology: Vol. 22. Systems and development* (pp. 167–209). Hillsdale, NJ: Erlbaum.

Patterson, G. R., & Cobb, J. A. (1971). A dyadic analysis of «aggressive» behaviors. In J. P. Hill (Ed.), *Minnesota symposia on child psychology* (Vol. 5, pp. 72–129). Minneapolis: University of Minnesota Press.

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys*. Eugene, OR: Castalia Publishing Company.

Pavot, W., Diener, E., Colvin, C. R., & Sandvik, E. (1991). Further validation of the Satisfaction with Life Scale: Evidence for the cross-method convergence of well-being measures. *Journal of Personality Assessment*, 57, 149–161.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research* (2nd ed.). New York: Holt, Rinehart & Winston.

Pedlow, R., Sanson, A., Prior, M., & Oberklaid, F. (1993). Stability of maternally reported temperament from infancy to 8 years. *Developmental Psychology*, 29, 998–1007.

Pence, A. R. (Ed.). (1988). *Ecological research with children and families: From concepts to methodology*. New York: Teachers College Press.

Perlmutter, M., Metzger, R., Miller, K., & Nezworski, T. (1980). Memory of historical events. *Experimental Aging Research*, 6, 47–60.

Perry, D. G., & Bussey, K. (1984). *Social development*. Englewood Cliffs, NJ: Prentice-Hall.

Perry, D. G., Bussey, K., & Freiberg, K. (1981). Impact of adults' appeals for sharing on the development of altruistic dispositions in children. *Journal of Experimental Child Psychology*, 32, 127–138.

Perry, D. G., & Perry, L. C. (1975). Observational subject's sex role behavior. *Journal of Personality and Social Psychology*, 31, 1083–1088.

Pfeiffer, E. (Ed.). (1975). *Multidimensional Functional Assessment: The OARS methodology*. Durham, NC: Center for the Study of Aging and Human Development.

Piaget, J. (1926). *The language and thought of the child*. New York: Harcourt Brace.

Piaget, J. (1929). *The child's conception of the world*. London: Routledge & Kegan Paul.

- Piaget, J. (1932). *The moral judgment of the child*. London: Roudedge & Kegan Paul.
- Piaget, J. (1951). *Play, dreams, and imitation in childhood*. New York: Norton.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.
- Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.
- Piaget, J., & Inhelder, B. (1956). *The child's conception of space*. London: Routledge & Kegan Paul.
- Piaget, J., & Inhelder, B. (1973). *Memory and intelligence*. New York: Basic Books.
- Piaget, J., & Inhelder, B. (1974). *The child's construction of quantities*. London: Routledge & Kegan Paul.
- Piaget, J., & Szeminska, A. (1952). *The child's conception of number*. Atlantic Highlands, NJ: Humanities.
- Piers, E., & Harris, D. (1969). *The Piers-Harris Children's Self-Concept Scale*. Nashville, TN: Counselor Recordings and Tests.
- Plomin, R. (1990). *Nature and nurture*. Belmont, CA: Wadsworth.
- Polyson, J., Levinson, M., & Miller, H. (1982). Writing styles: A survey of psychology journal editors. *American Psychologist*, 37, 335–338.
- Porges, S. W. (1979). Developmental designs for infancy research. In J. D. Osofsky (Ed.), *Handbook of infant development* (pp. 742–765). New York: Wiley.
- Posada, G., Gao, Y., Wu, F., Posada, R., Tascon, M., Schoelmerich, A., Sagi, A., Kondo-Ikemura, K., Haaland, W., & Synevaag, B. (1995). The securebase phenomenon across cultures: Children's behavior, mothers' preferences, and experts' concepts. In E. Waters, B. E. Vaughn, G. Posada, & K. Kondo-Ikemura (Eds.), *Caregiving, cultural, and cognitive perspectives on secure-base behavior and working models* (pp. 27–48). *Monographs of the Society for Research in Child Development*, 60 (2–3, Serial No. 244).
- Povinelh, D. J. (1995). The unduplicated self In P. Rochat (Ed.), *The self in early infancy* (pp. 161–192). Amsterdam: North-Holland-Elsevier.
- Pratt, C., & Bryant, P. E. (1990). Young children understand that looking leads to knowing (so long as they are looking through a single barrel). *Child Development*, 61, 973–982.
- Pressley, M., Levin, J. R., & Bryant, S. L. (1983). Memory strategy instruction during adolescence: When is explicit instruction needed? In M. Pressley & J. R. Levin (Eds.), *Cognitive strategy research: Psychological foundations* (pp. 25–49). New York: Springer-Verlag.
- Rabbitt, R. (1989). Inner-city decay? Age changes in structure and process in recall of familiar topographical information. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 284–299). New York: Cambridge University Press.
- Radke-Yarrow, M., & Zahn-Waxler, C. (1984). Roots, motive, and patterning in children's prosocial behavior. In E. Staub, D. Bar-Tal, J. Katylowski, & J. Raykowski (Eds.), *The development and maintenance of prosocial behavior* (pp. 81–99). New York: Plenum Press.

Reid, J. B. (1970). Reliability assessment of observation data: A possible methodological problem. *Child Development*, 41, 1143–1150.

Rest, J. R. (1979). *Development in judging moral issues*. Minneapolis: University of Minnesota Press.

Rheingold, H. L. (1982). Ethics as an integral part of research in child development. In R. Vasta (Ed.), *Strategies and techniques of child study* (pp. 305–325). New York: Academic Press.

Riccuiti, H. N., & Breitmayer, B. J. (1988). Observational assessments of infant temperament in the natural setting of the newborn nursery: Stability and relationship to perinatal status. *Merrill-Palmer Quarterly*, 34, 281–299.

Richardson, G. A., & McCluskey, K. A. (1983). Subject loss in infancy research: How biasing is it? *Infant Behavior and Development*, 6, 235–239.

Rogosa, D. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, 88, 245–258.

Rose, S. A., & Blank, M. (1974). The potency of context in children's cognition: An illustration through conservation. *Child Development*, 45, 499–502.

Rose, S. A., Schmidt, K., & Bridget, W. H. (1978). Changes in tactile responsivity during sleep in the human newborn infant. *Developmental Psychology*, 14, 163–172.

Rosenberg, M. (1979). *Conceiving the self*. New York: Basic Books.

Rosenberg, M. J. (1965). When dissonance fails: On eliminating evaluation apprehension from attitude measurement. *Journal of Personality and Social Psychology*, 1, 28–42.

Rosenthal, R. (1968). Experimenter expectancy and the reassuring nature of the null hypothesis decision procedure. *Psychological Bulletin Monograph Supplement*, 70, 30–47.

Rosenthal, R. (1976). *Experimenter effects in behavioral research* (enl. ed.). New York: Halsted Press.

Rosenthal, R. (1994a). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science*, 3, 176–179.

Rosenthal, R. (1994b). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.) *The handbook of research synthesis* (pp. 231–244). New York: Russell Sage Foundation.

Rosnow, R. L., & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist*, 44, 1276–1284.

Rosnow, R. L., & Rosenthal, R. (1995). «Some things you learn aren't so»: Cohen's paradox, Asch's paradigm, and the interpretation of interaction. *Psychological Science*, 6, 3–9.

Rosnow, R. L., & Rosnow, M. (1992). *Writing papers in psychology* (2nd ed.). Belmont, CA: Wadsworth.

Ross, J. B., & McLaughlin, M. M. (Eds.). (1949). *The portable medieval reader*. New York: Viking Press.

Rothbart, M. K. (1980, April). *Longitudinal home observation of infant temperament*. Paper presented at the International Conference on Infant Studies, New Haven, CT.

Rovee-Collier, C. K. (1987). Learning and memory in infancy. In J. D. Osofsky (Ed.), *Handbook of infant development* (2nd ed., pp. 98–148). New York: Wiley.

Rovee-Collier, C. K., & Bhatt, R. S. (1993). Evidence of long-term memory in infancy. In R. Vasta (Ed.), *Annals of child development* (Vol. 9, pp. 1–45). London: Kingsley.

Russell, A., Russell, G., & Midwinter, D. (1992). Observer influences on mothers and fathers: Self-reported influence during a home observation. *Merrill-Palmer Quarterly*, 38, 263–283.

Russell, J., Maudmer, N., Sharpe, S., & Tidswell, T. (1991). The «windows task» as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology*, 9, 331–349.

Rutherford, E., & Mussen, P. H. (1968). Generosity in nursery school boys. *Child Development*, 39, 755–765.

Sackett, G. P. (Ed.). (1978). *Observing behavior: Vol. 2. Data collection and analysis methods*. Baltimore. University Park Press.

Sackett, G. P., Ruppenthal, G. C., & Gluck, J. (1978). Introduction: An overview of methodological and statistical problems in observational research. In G. P. Sackett (Ed.), *Observing behavior. Vol. 2. Data collection and analysis methods* (pp. 1–14). Baltimore: University Park Press.

Salthouse, T. A. (1982). *Adult cognition: An experimental psychology of human aging*. New York: Springer-Verlag.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum.

Salthouse, T. A. (1994). The nature of the influence of speed on adult age differences in cognition. *Developmental Psychology*, 30, 240–259.

Sands, L. P., & Meredith, W. (1989). Effects of sensory and motor functioning on adult intellectual performance. *Journal of Gerontology: Psychological Sciences*, 44, 56–58.

Sawin, D. B., & Parke, R. D. (1980). Empathy and fear as mediators of resistance-to-deviation in children. *Merrill-Palmer Quarterly*, 26, 123–133.

Scarr, S., & Weinberg, R. A. (1983). The Minnesota Adoption Studies: Genetic differences and malleability. *Child Development*, 54, 260–267.

Schaffer, H. R., & Emerson, P. (1964). The development of social attachments in infancy. Monographs of the Society for Research in *Child Development*, 29 (3, Serial No. 94).

Schaie, K. W. (1958). Rigidity-flexibility and intelligence: A cross-sectional study of the adult life-span from 20 to 70. *Psychological Monographs*, 72 (9, Whole No. 462).

Schaie, K. W. (1993). The Seattle Longitudinal Studies of adult intelligence. *Current Directions in Psychological Science*, 2, 171–175.

Schaie, K. W. (1994). The course of adult intellectual development. *American Psychologist*, 49, 304–313.

Schaie, K. W. (1996a). Intellectual development in adulthood. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 266–286). San Diego: Academic Press.

Schaie, K. W. (1996b). *Intellectual development in adulthood: The Seattle Longitudinal Study*. Cambridge, Eng.: Cambridge University Press.

Schneider, B. A., Trehub, S. E., & Bull, D. (1980). High-frequency sensitivity in infants. *Science*, 207, 1003–1004.

Schonfield, A. E. D. (1969). *In search of early memories*. Paper presented at the International Congress of Gerontology, Washington, DC.

Schonfield, D., & Robertson, E. A. (1966). Memory storage and aging. *Canadian Journal of Psychology*, 20, 228–236.

Schwarz, J. C., Barton-Henry, M. L., & Pruzinsky, T. (1985). Assessing child-rearing behaviors: A comparison of ratings made by mother, father, child, and sibling on the CR-PB 1. *Child Development*, 56, 462–479.

Seitz, V. (1980). Statistical issues and comparative methods. In M. H. Bornstein (Ed.), *Comparative methods in psychology* (pp. 149–182). Hillsdale, NJ: Erlbaum.

Seitz V. (1984). Methodology. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental psychology: An advanced textbook* (pp. 37–79). Hillsdale, NJ: Erlbaum.

Shantz, C. U. (1975, April). *Communication skills and social-cognitive development*. Paper presented at the meeting of the Society for Research in Child Development, Denver, CO.

Sheikh, J. I., Hill, R. D., & Yesavage, J. (1986). Longterm efficacy of cognitive training for age-associated memory impairment: A six-month follow-up study. *Developmental Neuropsychology*, 2, 413–421.

Sieber, J. E. (1992). *Planning ethically responsible research: A guide for students and internal review boards*. Newbury Park, CA: Sage.

Siegel, M. (1991). *Knowing children: Experiments in conversation and cognition*. Hillsdale, NJ: Erlbaum.

Siegel, M., & Beattie, K. (1991). Where to look first for children's knowledge of false beliefs. *Cognition*, 38, 1–12.

Siegel, L. S. (1992). Infant motor, cognitive, and language behaviors as predictors of achievement at school age. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 7, pp. 227–237). Norwood, NJ: Ablex.

Siegel, S., & Castellan, N. J., Jr. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-HRI.

Siegler, I. C., & Botwinick, J. (1979). A long-term longitudinal study of intellectual ability of older adults: The matter of selective subject attrition. *Journal of Gerontology*, 34, 242–245.

Silverman, I. (1977). *The human subject in the psychological laboratory*. New York: Pergamon.

Slabach, E. H., Moffow, J., & Wachs, T. D. (1991). Questionnaire measurement of infant and child temperament: Current status and future directions. In S. Strelan &

A. Angleitner (Eds.), *Explorations in temperament: International perspectives on theory and measurement* (pp. 205–234). New York: Plenum Press.

Slaby, R. G., & Frey, K. S. (1975). Development of gender constancy and selective attention to same-sex models. *Child Development*, 46, 849–856.

Smetana, J. G. (1993). Understanding of social rules. In M. Bennett (Ed.), *The child as psychologist* (pp. 111–141). London: Guilford Press.

Smith, A. D. (1996). Memory. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 236–250). San Diego: Academic Press.

Smith, I. D. (1968). The effects of training procedures upon the acquisition of conservation of weight. *Child Development*, 39, 515–526.

Snarey, J. R. (1985). Cross-cultural universality of social-moral development: A critical review of Kohlbergian research. *Psychological Bulletin*, 97, 202–232.

Sockett, H. (1992). The moral aspects of the curriculum. In P. W. Jackson (Ed.), *Handbook of research on curriculum* (pp. 543–569). New York: Macmillan.

Sodian, B. (1994). Early deception and the conceptual continuity claim. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind* (pp. 385–401). Hillsdale, NJ: Erlbaum.

Spelke, E. S. (1991). Physical knowledge in infancy: Reflections on Piaget's theory. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays in biology and cognition* (pp. 133–169). Hillsdale, NJ: Erlbaum.

Spence, J. T., & Helmreich, R. L. (1978). *Masculinity and femininity: Their psychological dimensions, correlates, and antecedents*. Austin: University of Texas Press.

Spieth, W. (1965). Slowness of task performance and cardiovascular disease. In A. T. Welford & J. E. Biffen (Eds.), *Behavior, aging, and the nervous system* (pp. 366–400). Springfield, IL: Charles C. Thomas.

Sroufe, L. A. (1983). Infant-caregiver attachment and patterns of adaptation in preschool: The roots of maladaptation and competence. In M. Perlmutter (Ed.), *Minnesota symposia on child psychology: Vol. 16 Development and policy concerning children with special needs* (pp. 41–83). Hillsdale, NJ: Erlbaum.

Sroufe, L. A., Waters, E., & Matas, L. (1974). Contextual determinants of infant affective response. In M. Lewis & L. A. Rosenblum (Eds.), *The origins of fear* (pp. 49–72). New York: Wiley.

Stanley, B., & Sieber, J. E. (Eds.). (1991). *The ethics of research on children and adolescents*. Newbury Park CA: Sage.

Staub, E. (1971). The use of role playing and induction in children's learning of helping and sharing behavior. *Child Development*, 42, 805–816.

Steinberg, L., Elmen, J. D., & Mounts, N. S. (1989). Authoritative parenting, psychosocial maturity, and academic success among adolescents. *Child Development*, 60, 1424–1436.

Stem, M., & Karraker, K. H. (1989). Sex stereotyping of infants: A review of gender labeling studies. *Sex Roles*, 20, 501–522.

Stemberg, R. J. (1993). *The psychologist's companion* (3rd ed.). New York: Cambridge University Press.

Stevens, S. S. (1968). Measurement, statistics, and the schemapiric view. *Science*, 161, 849–856.

Stipek, D. J., Gralinski, J. H., & Kopp, C. B. (1990). Self-concept development in the toddler years. *Developmental Psychology*, 26, 972–977.

Storandt, M., & Hudson, W. (1975). Misuse of analysis of co-variance in aging research and some partial solutions. *Experimental Aging Research*, 1, 121–125.

Strayer, F. F., Verissimo, M., Vaughn, B. E., & Howes, C. (1995). A quantitative approach to the description and classification of primary social relationships. In E. Waters, B. E. Vaughn, G. Posada, & K. Kondo-Ikemura (Eds.), *Caregiving, cultural, and cognitive perspectives on secure-base behavior and working models* (pp. 49–70). *Monographs of the Society for Research in Child Development*, 60 (2–3, Serial No. 244).

Strunk, W., Jr., & White, E. B. (1979). *The elements of style* (3rd ed.). New York: Macmillan.

Sullivan, K., & Winner, E. (1993). Three-year-olds' understanding of mental states: The influence of trickery. *Journal of Experimental Child Psychology*, 56, 135–148.

Tannenbaum, A. S., & Cooke, R. A. (1977). Research involving children. In National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (Eds.), *Appendix to report and recommendations on research involving children* (pp. 111–129). Washington, DC: U.S. Government Printing Office.

Taplin, P. S., & Reid, J. B. (1973). Effects of instructional set and experimenter influence on observer reliability. *Child Development*, 44, 547–554.

Tatsuoka, M. (1993). Effect size. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 461–479). Hillsdale, NJ: Erlbaum.

Taylor, M., Ebsensen, B. M., & Bennett R. T. (1994). Children's understanding of knowledge acquisition: The tendency for children to report that they have always known what they have just learned. *Child Development*, 65, 1581–1604.

Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.

Thomas, A., Chess, S., & Birch, H. (1968). *Temperament and behavior disorders in children*. New York: New York University Press.

Thompson, R. A. (1990). Vulnerability in research: A developmental perspective on research risk. *Child Development*, 61, 1–16.

Thompson, R. A., & Hoffman, M. L. (1980). Empathy and the development of guilt in children. *Developmental Psychology*, 16, 155–156.

Thompson, S. K., & Bender, P. M. (1971). The priority of cues in sex discriminations by children and adults. *Developmental Psychology*, 5, 181–185.

Thomdike, R. L. (1933). The effect of the interval between test and retest on the constancy of IQ. *Journal of Educational Psychology*, 24, 543–549.

Thomdike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale* (4th ed.). Chicago: Riverside Publishing Company.

Thurstone, L. L., & Thurstone, T. G. (1962). *SRA Primary Mental Abilities*. Chicago: Science Research Associates.

Toussaint, N. A. (1974). An analysis of synchrony between concrete-operational tasks in terms of structure and performance demands. *Child Development*, 45, 992–1001.

Trials of war criminals before the Nuremberg military tribunals, U. S. vs. Karl Brandt (Vol. 2) (1949). Washington, DC: U. S. Government Printing Office.

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. New York: Cambridge University Press.

Turiel, E. (in press). Morality. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed.). New York: Wiley.

Turkheimer, E. (1991). Individual and group differences in adoption studies of IQ. *Psychological Bulletin*, 110, 392–405.

Tymchuk, A. J. (1991). Assent processes. In B. Stanley & J. E. Sieber (Eds.), *Social research on children and adolescents: Ethical issues* (pp. 128–139). Newbury Park, CA: Sage.

U. S. Bureau of the Census (1995). *Sixty-five plus in the United States*. Washington, DC: U.S. Government Printing Office.

Underwood, M. K., Coie, J. D., & Herbsman, C. R. (1992). Display rules for anger and aggression in school-age children. *Child Development*, 63, 366–380.

Unger, R. K., & Crawford, M. (1993). Commentary: Sex and gender-the troubled relationship between terms and concepts. *Psychological Science*, 4, 122–124.

Uttl, B., & Graf, P. (1993). Episodic spatial memory in adulthood. *Psychology and Aging*, 8, 257–273.

van Ijzendoorn, M. H. (1995). Adult attachment representations, parental responsiveness, and infant attachment: A meta-analysis on the predictive validity of the Adult Attachment Interview. *Psychologicalbulletin*, 117, 387–403.

van Uzendoorn, M. H., & Kroonenberg, P. M. (1988). Cross-cultural patterns of attachment: A metaanalysis of the Strange Situation. *Child Development*, 59, 147–156.

Vaughn, B. E., Stevenson-Hinde, J., Waters, E., Kotsaftis, A., Lefever, G. B., Shoul-dice, A., Trudel, M., & Belsky, J. (1992). Attachment security and temperament in infancy and early childhood: Some conceptual clarifications. *Developmental Psychology*, 28, 463–473.

Vaughn, B. E., & Waters, E. (1990). Attachment behavior at home and in the laboratory: Q-sort observations and Strange Situation classifications of one-year-olds. *Child Development*, 61, 1965–1973.

Verhaegen, P., Marcoen, A., & Goossens, L. (1992). Improving memory performance in the aged through mnemonic training: A meta-analytic study. *Psychology and Aging*, 7, 242–251.

- Voyat, G. E. (1982). *Piaget systematized*. Hillsdale, NJ: Erlbaum.
- Walk, R. D., & Gibson, E. J. (1961). A comparative and analytical study of visual depth perception. *Psychological Monographs*, 75 (15, Whole No. 519).
- Walker, L. J. (1988). The development of moral reasoning. In R. Vasta (Ed.), *Annals of child development* (Vol. 5, pp. 33–78). Greenwich, CT: JAI Press.
- Walker, L. J. (1989). A longitudinal study of moral reasoning. *Child Development*, 60, 157–166.
- Walker, L. J., devries, B., & Trevethan, S. D. (1987). Moral stages and moral orientations in real-life and hypothetical dilemmas. *Child Development*, 58, 842–848.
- Walker, L. J., & Taylor, J. H. (1991). Stage transitions in moral reasoning: A longitudinal study of developmental processes. *Developmental Psychology*, 27, 330–337.
- Wallach, L., Wall, A. J., & Anderson, L. (1967). Number conservation: The role of reversibility, addition-subtraction, and misleading perceptual cues. *Child Development*, 38, 425–442.
- Waters, E. (1995). The Attachment Q-Set (version 3.0). In E. Waters, B. E. Vaughn, G. Posada, & K. Kondo-Ikemura (Eds.), *Caregiving, cultural, and cognitive perspectives on secure-base behavior and working models* (pp. 234–246). *Monographs of the Society for Research in Child Development*, 60 (2–3, Serial No. 244).
- Waters, E., & Deane, K. E. (1985). Defining and assessing individual differences in attachment relationship: Q-methodology and the organization of behavior in infancy and early childhood. In I. Bretherton & E. Waters (Eds.), *Growing points of attachment theory and research* (pp. 41–65). *Monographs of the Society for Research in Child Development*, 50 (1–2, Serial No. 209).
- Webb, E. J., Campbell, D. T., Schwartz, R. D., Sechrest, L., & Grow, J. B. (1981). *Nonreactive measures in the social sciences* (2nd ed.). Boston: Houghton Mifflin.
- Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4th ed.). Baltimore: Williams & Wilkins.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. New York: The Psychological Corporation.
- Wechsler, D. (1989). *Wechsler Preschool and Primary Scale of Intelligence-Revised*. New York: The Psychological Corporation.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children-Third Edition*. New York: The Psychological Corporation.
- Weizmann, F., Cohen, L. B., & Pratt, R. J. (1971). Novelty, familiarity, and the development of infant attention. *Developmental Psychology*, 4, 149–154.
- Wellman, H. M., Ritter, K., & Flavell, J. H. (1975). Deliberate memory behavior in the delayed reactions of very young children. *Developmental Psychology*, 11, 780–787.
- West, R. L. (1995). Compensatory strategies for age-associated memory impairment. In A. D. Baddely, B. A. Wilson, & F. N. Watts (Eds.), *Handbook of memory disorders* (pp. 481–500). New York: Wiley.
- West, R. L., & Sinnott, J. D. (Eds.). (1992). *Everyday memory and aging*. New York: Springer-Verlag.

West, R. L., & Walton, M. (1985, March). *Practical memory functioning in the elderly*. Paper presented at the meeting of the Second National Forum on Research in Aging, Lincoln, NE.

Wetherford, M. J., & Cohen, L. B. (1973). Developmental changes in infant visual preferences for novelty and familiarity. *Child Development*, 44, 416–424.

Willis, S. L. (1996). Everyday problem solving. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 287–307). San Diego: Academic Press.

Willis, S. L., & Schaie, K. W. (1994). Assessing everyday competence in the elderly. In C. B. Fisher & R. M. Lemer (Eds.), *Applied developmental psychology* (pp. 339–372). New York: McGraw-Hill.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103–128.

Windle, M., & Lemer, R. M. (1986). Reassessing the dimensions of temperament individuality across the life span: The revised Dimensions of Temperament Survey (DTS-R). *Journal of Adolescent Research*, 1, 213–230.

Winterling, D., Crook, T. H., Salama, M., & Gobert, J. (1986). A self-rating scale for assessing memory loss. In A. Bes, J. Cahn, S. Hoyer, J. P. Marc-Vergnes, & H. M. Wisniewski (Eds.), *Senile dementias: Early detection* (pp. 482–486). London: John Libbey Eurotext.

World Medical Association (1964). Declaration of Helsinki. In H. K. Beecher (Ed.), *Research and the individual: Human studies* (pp. 277–278). Boston: Little, Brown.

Worobey, J., & Brazelton, T. B. (1990). Newborn assessment and support for parenting: The Neonatal Behavioral Assessment Scale. In E. D. Gibbs & D. M. Teti (Eds.), *Interdisciplinary assessment of infants* (pp. 33–44). Baltimore: Paul H. Brookes Publishing Co.

Yarrow, M. R., Campbell, J. D., & Burton, R. V. (1968). *Child rearing: An inquiry into research and methods*. San Francisco: Jossey-Bass.

Yarrow, M. R., Campbell, J. D., & Burton, R. V. (1970). Recollections of childhood: A study of the retrospective method. *Monographs of the Society for Research in Child Development*, 35, (5, Serial No. 138).

Yarrow, M. R., & Waxler, C. Z. (1976). Dimensions and correlates of prosocial behavior in young children. *Child Development*, 47, 118–125.

Yarrow, M. R., & Waxler, C. Z. (1979). Observing interaction: A confrontation with methodology. In R. B. Cairns (Ed.), *The analysis of social interactions: Methods, issues, and illustrations* (pp. 37–65). Hillsdale, NJ: Erlbaum.

Yonas, A. (1981). Infants' responses to optical information for collision. In R. N. Aslin, J. R. Alberts, & M. R. Peterson (Eds.), *Development of perception: Psychobiological perspectives: Vol. 2. The visual system* (pp. 313–334). New York: Academic Press.

Zahn-Waxler, C., & Radke-Yarrow, M. (1982). The development of altruism: Alternative research strategies. In N. Eisenberg (Ed.), *The development prosocial behavior* (pp. 109–137). San Diego: Academic Press.

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, 28, 126–136.

Zelnicker, T., Oppenheimer, L., & Renan, A. (1975). Effect of dimensional salience and salience of variability on problem solving: A developmental study. *Developmental Psychology*, 11, 334–341.

Zigler, E., & Finn-Stevenson, M. (1992). Applied developmental psychology. In M. H. Bornstein & M. E. Lamb (Eds.), *Advanced developmental psychology* (3rd ed., pp. 677–729). Hillsdale, NJ: Erlbaum.

Zigler, E. F., & Muenchow, S. (1992). *Head Start: The inside story of America's most successful educational experiment*. New York: Basic Books.

Zigler, E. F., & Styfco, S. (Eds.). (1993). *Head Start and beyond: A national plan for extended childhood intervention*. New Haven, CT: Yale University Press.

Zimmerman, B. J., & Rosenthal, T. L. (1974). Conserving and retaining equalities and inequalities through observation and correction. *Developmental Psychology*, 10, 260–268.

Глоссарий

1. **Валидность (validity)** — обоснованность выводов, получаемых в исследовании.
2. **Валидность статистического вывода (statistical conclusion validity)** — обоснованность статистических выводов, полученных при анализе данных.
3. **Валидность теста (test validity)** — точность, с которой тест измеряет то, что он призван измерять.
4. **Величина эффекта (effect size)** — показатель силы значимой связи между независимыми и зависимыми переменными.
5. **Взаимодействие (interaction)** — ситуация, когда влияние на зависимую переменную одной независимой переменной изменяется вместе с уровнем другой независимой переменной.
6. **Видимость—реальность (appearance/reality distinction)** — осознание различий между тем, как выглядит объект, и тем, что он есть в действительности.
7. **Визуальное самоузнавание (visual self-recognition)** — способность ребенка узнавать самого себя, обычно оценивается по узнаванию своего отражения в зеркале.
8. **«Визуальный обрыв» (visual cliff)** — метод изучения восприятия глубины в период младенчества; с его помощью измеряется реакция младенца на кажущийся резкий перепад высоты под стеклянной поверхностью.
9. **Включенное наблюдение (participant observation)** — одна из форм обсервационного исследования, в которой наблюдение ведет лицо, являющееся естественной частью соответствующей среды.
10. **Влияние наблюдателя (observer influence)** — искажающее влияние присутствия наблюдателя на наблюдаемое поведение.
11. **Внешняя валидность (external validity)** — обоснованность генерализации результатов исследования.
12. **Внутренняя валидность (internal validity)** — обоснованность выводов о причинно-следственных связях между исследуемыми переменными.
13. **Внутрисубъектный план (within-subject design)** — экспериментальный план, при котором одни и те же испытуемые ставятся в экспериментальные условия, соответствующие разным уровням независимой переменной.
14. **Воспроизведение (recall)** — восстановление в памяти воспринятых ранее, но отсутствующих в данный момент стимулов.
15. **Выборка (sample)** — подгруппа представителей популяции, отбираемая в целях исследования.
16. **Выборочный анализ поведенческих событий (event sampling)** — одна из форм наблюдения, при которой наблюдатель производит регистрацию только в период, когда имеет место интересующее его поведение.
17. **Габитуация (habituation)** — угасание ориентировочной реакции при неоднократном воздействии стимула.
18. **Главный эффект (main effect)** — результат прямого влияния независимой переменной на зависимую, то есть влияния, не зависящего от влияния других независимых переменных.

19. **Дебрифинг (debriefing)** — этический принцип, согласно которому в случае неполного информирования или обмана после эксперимента испытуемому необходимо разъяснить истинные цели исследования.
20. **Дегабитуация (dishabituation)** — восстановление ориентировочной реакции в ответ на изменение стимула, по отношению к которому возникло привыкание.
21. **Дескриптивный статистический показатель (descriptive statistic)** — статистический показатель, предназначенный для структурирования и обобщения данных.
22. **Детская биография (baby biography)** — метод исследования, при котором родитель ведет наблюдение за развитием собственного ребенка.
23. **Дисперсия (variance)** — среднее квадратичных отклонений от группового среднего — мера изменчивости.
24. **Дисперсия общности методов (shared methods variance)** — сходство двух или более совокупностей показателей, обусловленное сходством методов, с помощью которых они были получены.
25. **Диффузия (diffusion)** — непредусмотренный перенос эффекта экспериментального воздействия с экспериментальной группы на контрольную.
26. **Долговременная память (remote memory)** — память на события далекого прошлого.
27. **Дрейф наблюдателя (observer drift)** — снижение уровня надежности наблюдений с ослаблением контроля за наблюдателями.
28. **Заблуждение (false belief)** — вера в то, что не соответствует действительности.
29. **Зависимая переменная (dependant variable)** — переменная, которую измеряет исследователь при изменении независимой переменной.
30. **Избирательный отсев (selective dropout)** — систематическое выбывание испытуемых в ходе исследования.
31. **Изменчивость (variability)** — степень разброса показателей в некой совокупности.
32. **Исследование «вслепую» (blinding)** — сокрытие от лиц, проводящих тестирование или наблюдение, информации, способной исказить результаты.
33. **Институциональный наблюдательный совет (institutional review board)** — организация, в обязанности которой входит оценка этичности предлагаемых исследований до начала сбора данных.
34. **Инструментация (instrumentation)** — незапланированные изменения, происходящие с экспериментаторами, наблюдателями или измерительными инструментами в ходе исследования.
35. **Интервальная шкала (interval scale)** — одна из шкал измерения, в которой деления представляют собой равные количественные интервалы.
36. **Информированное согласие (informed consent)** — согласие участвовать в исследовании, подкрепленное осведомленностью о сути и последствиях участия.
37. **История (history)** — потенциально значимые события, происходящие в период между измерениями и оказывающие воздействие наряду с изучаемой независимой переменной.

38. **Исходный уровень (baseline)** — характеристики поведения до экспериментального вмешательства.
39. **Клинический метод (clinical method)** — гибкий, полустандартизованный метод постановки перед детьми задач и вопросов, который использовал Пиаже.
40. **Когорта (cohort)** — группа людей, имеющих некий общий опыт и некие общие характеристики — чаще всего определяется по году рождения.
41. **Коммуникация между испытуемыми (between-subject communication)** — незапланированное общение между испытуемыми, способное привести к искажению результатов исследования.
42. **Конвергентные операции (converging operations)** — использование разнообразных методов для изучения определенной проблематики.
43. **Конструктивная память (constructive memory)** — влияние системы знаний на мнемические процессы.
44. **Конструктная валидность (исследования) (construct validity (of study))** — точность теоретической интерпретации результатов исследования.
45. **Конструктная валидность (теста) (construct validity (of test))** — одна из форм валидности теста, которая зависит от того, в какой мере разного рода данные подтверждают обоснованность теоретической интерпретации измеряемого атрибута.
46. **Контрбалансировка (counterbalancing)** — во внутрисубъектных планах исследования изменение порядка следования заданий или экспериментальных условий.
47. **Конфиденциальность (confidentiality)** — этический принцип, согласно которому информация, получаемая в исследовании, должна быть недоступна в той мере, в какой она может навредить испытуемым или смутить их.
48. **Корреляционное исследование (correlational research)** — одна из форм исследования, в которой не контролируются независимые переменные, а анализируется связь между двумя или более зависимыми переменными — другое название *неэкспериментальное исследование*.
49. **Корреляционный статистический показатель (correlation statistic)** — индекс силы и направления связи двух переменных.
50. **Кристаллизованный интеллект (crystallized intelligence)** — система знаний, накопленных через опыт.
51. **Критериальная валидность (criterion validity)** — одна из форм валидности теста, которая зависит от корреляции показателей теста с определенными внешними критериями измеряемого атрибута.
52. **Лонгитюдный план (longitudinal design)** — план исследования, предназначенный для возрастного сравнения, в котором в течение определенного времени в разном возрасте обследуются одни и те же испытуемые.
53. **Макросистема (macrosystem)** — культура или субкультура, в которой живет человек — четвертый из четырех уровней контекстов в теории экологических систем Бронфенбреннера.

54. **Макросистема наблюдения (molar observational system)** — система наблюдения, в которой при оценке поведения используются глобальные, интерпретационные категории.
55. **Матрица «свойства × методы» (multitrait-multimethod matrix)** — метод оценки конструктивной валидности тестов — связан с анализом матрицы корреляций, составленной по результатам измерения по меньшей мере двух конструктов по крайней мере двумя методами.
56. **Медиана (median)** — середина в распределении совокупности показателей.
57. **Межсубъектный план (between-subject design)** — экспериментальный план исследования, при котором разных испытуемых ставят в условия, соответствующие разным уровням независимой переменной.
58. **Мезосистема (mesosystem)** — взаимосвязи между микросистемами среды — второй из четырех уровней контекста в теории экологических систем Бронфенбреннера.
59. **Метод временных срезов (time sampling)** — одна из форм обсервативной оценки, при которой поведение регистрируется в течение определенных (обычно непродолжительных) интервалов времени.
60. **Метод парциальной корреляции (partial correlation technique)** — процедура статистического исключения влияния потенциально значимого третьего фактора на корреляцию двух других переменных.
61. **Метод предпочтений (preference method)** — метод изучения зрительных возможностей в период младенчества — с его помощью оценивается внимание младенца к двум одновременно предъявляемым стимулам.
62. **Методики самоотчета (self-report measure)** — методики, при использовании которых о характеристиках испытуемых судят со слов самих испытуемых.
63. **Микросистема (microsystem)** — ближайшее окружение человека — первый из четырех уровней контекста в теории экологических систем Бронфенбреннера.
64. **Микросистема наблюдения (molecular observational system)** — система наблюдения, в которой оцениваются специфические детали поведения.
65. **Мнемонические стратегии (mnemonic strategies)** — приемы (например, повторение, упорядочение), используемые для облегчения запоминания.
66. **Мода (mode)** — наиболее часто встречающееся значение в совокупности показателей — мера центральной тенденции.
67. **Монометодическое искажение (mono-method bias)** — искажение, являющееся следствием использования только одного метода для выявления отношений между независимой и зависимой переменными.
68. **Монооперациональное искажение (mono-operation bias)** — искажение, являющееся следствием использования только одной операционализации независимой или зависимой переменной.
69. **Мощность (power)** — вероятность того, что статистический критерий исключает ложную нуль-гипотезу.
70. **Надежность (reliability)** — согласованность или воспроизводимость результатов измерения.

71. **Надежность внутренней согласованности (internal consistency reliability)** — одна из форм надежности, основанная на согласованности ответов, или реакция на разные компоненты теста.
72. **Нарративная запись (narrative record)** — одна из форм обсервативного измерения, при которой наблюдатель непрерывно описывает поведение испытуемого.
73. **Независимая оценка (independent review)** — этический принцип, согласно которому все исследования с испытуемыми людьми должны подвергаться независимой проверке на соответствие этическим нормам.
74. **Независимая переменная (independent variable)** — переменная, которую исследователь контролирует посредством манипуляций или отбора с целью оценки ее влияния на зависимую переменную.
75. **«Незнакомая ситуация» (Strange Situation)** — структурированный лабораторный метод оценки привязанности — с его помощью измеряется реакция младенца на ряд эпизодов разлучения и воссоединения с матерью.
76. **Необъективность наблюдателя (observer bias)** — тенденция воспринимать и регистрировать поведение предвзято, в соответствии с имеющимися ожиданиями.
77. **Непараметрический критерий (nonparametric test)** — статистический критерий, валидность которого не зависит от распределения показателей в изучаемой популяции.
78. **Неполное информирование (incomplete disclosure)** — предоставление неполной информации об исследовании.
79. **Номинальная шкала (nominal scale)** — шкала измерения, градация которой соответствует наименованиям, а не количественным характеристикам.
80. **Нормальное распределение (normal distribution)** — колоколообразное распределение показателей, в котором среднее, медиана и мода совпадают.
81. **Обман (deception)** — преднамеренное введение в заблуждение относительно исследования.
82. **Обусловленный поворот головы (conditioned head turning)** — метод изучения перцептивных возможностей в период младенчества — с его помощью оценивается способность младенца распознавать и различать стимулы по наличию или отсутствию реакции в форме обусловленного поворота головы при появлении стимула.
83. **Операциональное определение (operational definition)** — определение переменной с точки зрения операций, используемых для ее создания или измерения.
84. **Ориентировочная реакция (orienting response)** — естественная реакция внимания на новые стимулы.
85. **Отбор (sampling)** — создание выборки испытуемых из некой популяции.
86. **Отказ в благотворном воздействии (withholding treatment)** — решение не включать кого-либо в экспериментальную группу, которая подвергается потенциально благотворному воздействию.
87. **Оценочная шкала поведения новорожденного (Neonatal Behavioral Assessment Scale)** — стандартизованная процедура оценки статуса разви-

- тия новорожденных — с ее помощью измеряются как рефлексy, так и ранние формы адаптивного поведения.
88. **Ошибка второго рода (Type 2 error)** — ошибочное принятие ложной нуль-гипотезы.
89. **Ошибка первого рода (Type 1 error)** — ошибочный отказ от истинной нуль-гипотезы.
90. **Параметрический критерий (parametric test)** — статистический критерий, валидность которого зависит от распределения показателей изучаемой популяции.
91. **Первичная дисперсия (primary variance)** — различия, объясняемые действием независимых переменных исследования.
92. **Перекрестная корреляционная группировка с лагом (cross-lagged panel correlation)** — в лонгитюдном исследовании, анализ отношений между двумя переменными в рамках одного периода времени и в рамках разных периодов времени с целью выявления направления причинного воздействия в корреляционной паре.
93. **Персональное влияние (personal agency)** — понимание ребенком того, что он может являться причиной некоторых событий.
94. **План временных серий (time-series design)** — одна из форм внутрисубъектного плана, при которой производится неоднократное измерение зависимой переменной при наличии и отсутствии экспериментального воздействия.
95. **План поперечных срезов (cross-sectional design)** — план исследования, в котором одновременно изучаются и сравниваются испытуемые разных возрастов.
96. **План с временным лагом (time-lag design)** — план исследования, при котором в разное время обследуются испытуемые одного возраста, но принадлежащие к разным когортам.
97. **План уравненных групп (matched groups design)** — одна из форм межсубъектного плана, в которой до причисления испытуемых к разным экспериментальным условиям производится отбор по потенциально значимым характеристикам так, чтобы между испытуемыми было соответствие.
98. **Понятие об объекте (object concept)** — осознание того, что объект существует независимо от перцептивного контакта с ним.
99. **Популяция (population)** — определенная группа людей, совокупность результатов измерения или событий.
100. **Порядковая шкала (ordinal scale)** — шкала измерения, в которой деления соответствуют рангам количественных величин.
101. **Последовательный анализ (sequential analysis)** — процедура выявления причинно-следственных отношений в обсервативном исследовании, заключающаяся в установлении временных отношений между действиями двух или более взаимодействующих людей.
102. **Последовательный план (sequential design)** — план исследования, при котором с целью оценки влияния факторов возраста, когорты и времени

измерения сочетаются лонгитюдный метод, метод поперечных срезов и метод с временным лагом.

103. **Постоянство пола (gender constancy)** — понимание того, что пол — неизменное качество.
104. **Привязанность (attachment)** — эмоциональная связь между младенцем и заботящимся о нем лицом, формирующаяся в период младенчества.
105. **«Привязанность» Q-набор (Attachment Q-set)** — методика рейтинга, предназначенная для оценки привязанности и основанная на суждениях о том, насколько характеристики, свидетельствующие о привязанности, типичны для младенца
106. **Просоциальное поведение (prosocial behavior)** — поведенческое проявление нравственности — социальное одобряемое поведение, например оказание помощи, акты дарения.
107. **Реактивность (reactivity)** — результат непредусмотренного влияния экспериментальных мероприятий на поведение испытуемого.
108. **Регистрация движений глаз (eye-movement recording)** — метод изучения зрительных возможностей в период младенчества — с его помощью измеряется последовательность фиксаций взгляда младенца при исследовании им визуального стимула.
109. **Регрессия к среднему (regression toward the mean)** — тенденция приближения изначально крайне высоких или крайне низких показателей к групповому среднему значению.
110. **Репрезентационное изменение (representational change)** — осознание возможности изменения собственных представлений.
111. **Ретестовая надежность (test-retest reliability)** — одна из форм надежности, зависящая от соответствия результатов неоднократного предъявления одного и того же теста.
112. **Систематическая ошибка при отборе (selection bias)** — причисление к сравниваемым группам изначально неравноценных испытуемых.
113. **Сложение классов (class inclusion)** — понимание того, что подкласс не может быть больше класса, в который он входит.
114. **Случайное причисление (random assignment)** — процедура причисления испытуемых к экспериментальным условиям, в которой шансы быть причисленным к тем или иным условиям равны.
115. **Случайный отбор (random sampling)** — процедура отбора испытуемых, при которой все представители изучаемой популяции имеют равные шансы войти в выборку.
116. **Смешение (confounding)** — непредусмотренная связь между двумя потенциально значимыми переменными.
117. **Содержательная валидность (content validity)** — одна из форм валидности теста, которая зависит от адекватности отображения тестом измеряемого атрибута.
118. **Созревание (maturation)** — естественные изменения испытуемого как функция от времени, в течение которого проводится исследование.

119. **Состояние (state)** — уровень психической и поведенческой активности, колеблющийся от глубокого сна до напряженного плача.
120. **Сохранение (conservation)** — осознание того, что количественные характеристики объекта не изменяются при изменении его внешнего вида.
121. **Среднее (mean)** — среднее арифметическое в совокупности показателей — мера центральной тенденции.
122. **Стандартизация (standardization)** — создание единообразия всех аспектов экспериментальной процедуры в определенных условиях для всех испытуемых.
123. **Стандартное отклонение (standard deviation)** — квадратный корень из показателя дисперсии — мера изменчивости.
124. **Статистический показатель, выводимый логическим путем (inferential statistic)** — статистический показатель, используемый для определения уровня значимости различий между сравниваемыми группами.
125. **Субъектная переменная (subject variable)** — независимая переменная, уровень которой отражает закрепленные характеристики испытуемого, его неотъемлемые качества, такие как пол или возраст.
126. **Сценарий (script)** — порядок и структура определенных событий, зафиксированные в представлениях.
127. **Текущий интеллект (fluid intelligence)** — базовые навыки решения задач, в целом не зависящие от опыта.
128. **Теория сознания (theory of mind)** — суждения и представления о психическом мире.
129. **Терминальный спад (terminal drop)** — снижение умственных возможностей в период, непосредственно предшествующий смерти.
130. **Тестирование (testing)** — влияние тестирования на результаты последующего тестирования.
131. **Транзитивность (transitivity)** — способность логическим путем прийти к заключению об отношениях между некоторыми величинами; например, если $A > B$, а $B > C$, то $A > C$.
132. **Узнавание (recognition)** — осознание того, что воспринимаемый в данный момент стимул уже воспринимался ранее.
133. **Уровни (levels)** — значения, которые в исследовании принимает независимая переменная.
134. **Установка на определенный ответ (response set)** — зафиксированная, искаженная форма реагирования, не зависящая от содержания задания.
135. **Устойчивость (robustness)** — мера независимости статистического критерия от нарушения лежащих в его основе предположений.
136. **Фактор (factor)** — синоним независимой переменной.
137. **Характеристики выдвигаемых требований (demand characteristics)** — особенности экспериментальных условий, по которым испытуемые судят о том, каковы ожидания в отношении их реакции.
138. **Центральная тенденция (central tendency)** — показатель, соответствующий преобладающему паттерну реакции в выборке.

139. **Шкала отношений (ratio scale)** — шкала измерения, деления которой соответствуют равным количественным интервалам, а сама шкала включает нулевой уровень.
140. **Эквивалентность измерения (measurement equivalence)** — сопоставимость процедур и методов измерения, используемых в сравниваемых группах.
141. **Экзосистема (exosystem)** — социальные системы, оказывающие влияние на человека, в которых он, однако, не принимает непосредственного участия — третий из четырех уровней контекста в теории экологических систем Бронфенбреннера.
142. **Эффект верхнего предела (ceiling effect)** — получение показателей зависимой переменной, соответствующих максимально возможным значениям или приближающихся к ним.
143. **Эффект нижнего предела (floor effect)** — получение показателей зависимой переменной, соответствующих минимально возможному значению или приближающихся к нему.
144. **Эффект переноса (carry-over effect)** — результат влияния на реакцию испытуемого при выполнении определенного задания или на его поведение в определенных экспериментальных условиях предшествующих заданий или экспериментальных условий.
145. **Эффект последовательности (order effect)** — систематическое изменение реакции испытуемого от начала к концу эксперимента.
146. **Эффект предубеждений экспериментатора (experimenter bias effect)** — искажающее влияние ожиданий экспериментатора на результаты исследования.

УВАЖАЕМЫЕ ГОСПОДА!

**ИЗДАТЕЛЬСКИЙ ДОМ «ПИТЕР» ПРИГЛАШАЕТ ВАС К ВЗАИМОВЫГОДНОМУ СОТРУДНИЧЕСТВУ.
ЗА ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИЕЙ ОБРАЩАЙТЕСЬ В НАШИ ПРЕДСТАВИТЕЛЬСТВА,
КОТОРЫЕ НАХОДЯТСЯ ПО СЛЕДУЮЩИМ АДРЕСАМ:**

Россия, г. Москва

Представительство издательства «Питер», м. «Калужская»,
ул. Булгерова, д. 170, оф. 207 и 240, тел./факс (095) 777-54-67.
E-mail: sales@piter.msk.ru

Россия, г. Санкт-Петербург

Представительство издательства «Питер», м. «Электросила»,
ул. Благодатная, д. 67, тел.: (812) 327-93-37, 294-54-65.
E-mail: sales@piter.com

Украина, г. Харьков

Представительство издательства «Питер», тел. (0572) 14-96-09,
факс: (0572) 28-20-04, 28-20-05. Почтовый адрес: 61093,
г. Харьков, а/я 9130.
E-mail: piter@tender.kharkov.ua

Украина, г. Киев

Филиал Харьковского представительства издательства «Питер»,
тел./факс: (044) 490-35-68, 490-35-69. Адрес для писем: 04116,
г. Киев-116, а/я 2. Фактический адрес: 04073, г. Киев,
пр. Красных Казаков, д. 6, корп. 1.
E-mail: office@piter-press.kiev.ua

Беларусь, г. Минск

Представительство издательства «Питер», тел./факс (37517)
239-36-56. Почтовый адрес: 220100, г. Минск, ул. Куйбышева, 75.
ООО «Питер М», книжный магазин «Эврика».
E-mail: piterbel@tut.by

**КАЖДОЕ ИЗ ЭТИХ ПРЕДСТАВИТЕЛЬСТВ
РАБОТАЕТ ПО ЕДИНОМУ СТАНДАРТУ
ИЗДАТЕЛЬСКОГО ДОМА «ПИТЕР».**

**КНИГИ ИЗДАТЕЛЬСКОГО ДОМА «ПИТЕР» ВЫ МОЖЕТЕ ПРИОБРЕСТИ ОПТОМ И В РОЗНИЦУ
У НАШИХ РЕГИОНАЛЬНЫХ ПАРТНЕРОВ.**

Башкортостан

Уфа, «Азия», ул. Зенцова, д. 70 (оптовая продажа), маг. «Оазис»,
ул. Чернышевского, д. 88, тел./факс (3472) 50-39-00.
E-mail: asiaufa@ufanet.ru

Дальний Восток

Владивосток, «Приморский Торговый дом книги»,
тел./факс (4232) 23-82-12. Почтовый адрес: 690091, Владивосток,
ул. Светланская, д. 43.
E-mail: bookbase@mail.primorye.ru

Хабаровск, «Мирс», тел. (4212) 30-54-47, факс 22-73-30.

Почтовый адрес: 680000, Хабаровск, ул. Ким-Ю-Чена, д. 21.
E-mail: postmaster@bookmirs.khv.ru

Хабаровск, «Книжный мир», тел. (4212) 32-85-51, факс 32-82-50.

Почтовый адрес: 680000, Хабаровск, ул. Карла Маркса, д. 37.
E-mail: postmaster@worldbooks.knt.ru

Европейские регионы России

Архангельск, «Дом книги», тел. (8182) 65-41-34, факс 65-41-34.
Почтовый адрес: 163061, пл. Ленина, д. 3.
E-mail: book@atnet.ru

Калининград, «Вестер», тел./факс (0112) 21-56-28, 21-62-07.

Почтовый адрес: 236040, Калининград, ул. Победы, д. 6.
Магазин «Книги & книжечки».
E-mail: nshibkova@vester.ru; www.vester.ru

Ростов-на-Дону, ПБОЮЛ Остроменский, пр. Соколова, д. 73,
тел./факс (8632) 32-18-20. E-mail: ostrom@don.sitek.net

Северный Кавказ

Ессентуки, «Россы», ул. Октябрьская, 424,
тел./факс (87934) 6-93-09. E-mail: rossy@kmmw.ru

Сибирь

Иркутск, «Продавить», тел. (3952) 59-13-70, факс 51-30-70.
Почтовый адрес: 664031,

Иркутск, ул. Байкальская, д. 172, а/я 1397.

E-mail: prodalit@irk.ru; http://www.prodalit.irk.ru

Иркутск, «Антей-книга», тел./факс (3952) 33-42-47. Почтовый адрес: 664003, Иркутск, ул. Карла Маркса, д. 20.

E-mail: antey@irk.ru

Красноярск, «Книжный мир», тел./факс (3912) 27-39-71. Почтовый адрес: 660049, Красноярск, пр. Мира, д. 86.

E-mail: book-world@public.krasnet.ru

Нижневартовск, «Дом книги», тел. (3466) 23-27-14, факс 23-59-50.

Почтовый адрес: 628606, Нижневартовск, пр. Победы, д. 12.
E-mail: book@nvartovsk.wsnet.ru

Новосибирск, «Топ-книга», тел. (3832) 36-10-26, факс 36-10-27.

Почтовый адрес: 630117, Новосибирск, а/я 560.
E-mail: office@top-kniga.ru; http://www.top-kniga.ru

Тюмень, «Друг», тел./факс (3452) 21-34-39, 21-34-82. Почтовый адрес: 625019, ул. Республики, д. 211.

E-mail: drug@tyumen.ru

Тюмень, «Фолиант», тел. (3452) 27-36-06, факс 27-36-11. Почтовый адрес: 625039, Тюмень, ул. Харьковская, д. 83а.

E-mail: foliant@tyumen.ru

Татарстан

Казань, «Таис», тел. (8432) 72-34-55, факс 72-27-82. Почтовый
адрес: 420073, Казань, ул. Гвардейская, д. 9а.
E-mail: tais@bancorp.ru

Урал

Екатеринбург, магазин № 14, ул. Челюскинцев, д. 23, тел./факс
(3432) 53-24-90. E-mail: gvardia@mail.ur.ru

Екатеринбург, «Валео-книга», ул. Ключевская, д. 5, тел. (3432) 42-07-75, факс 42-56-00.

E-mail: valeo@etel.ru

Скотт А. Миллер — профессор психологии Университета Флориды, доктор философии, член Американской психологической ассоциации. Сферы его научно-исследовательских интересов: детская психология, когнитивные аспекты развития.

Трудность проведения исследования в области психологии развития — сквозная тема предлагаемой книги. Автор рассматривает ряд основных законов и принципов, которые необходимо учитывать при осуществлении собственных и оценке сторонних исследований. Это издание преследует три основные цели. Первая из них — помочь выработать навыки, необходимые для того, чтобы качественно изучать различные феномены в психологии развития, в связи с чем представлены разного рода принципы и инструкции, часть которых касаются непосредственно вопросов развития, а другие относятся к психологии в целом. Вторая цель заключается в ознакомлении читателей с важнейшими направлениями научной работы в данной области. Третья цель — сформировать не только навыки критической оценки любого исследования, но и умение на его основе сделать адекватные выводы.

Эта книга адресована всем желающим углубить свои знания о проблемах и методах исследований в психологии развития: студентам, аспирантам, преподавателям факультетов психологии, психологам и психотерапевтам.

КНИГИ ИЗДАТЕЛЬСТВА «ПИТЕР»



Спрашивайте в книжных магазинах
или заказывайте наложенным платежом:

197198, Санкт-Петербург, а/я 619
e-mail: sales@piter.com
для жителей России

61093, Харьков-93, а/я 9130,
e-mail: piter@tender.kharkov.ua
для жителей Украины



ПИТЕР
WWW.PITER.COM

Посетите наш web-магазин:
<http://www.piter.com>

ISBN 5-318-00472-5



9 785318 004728